Heidelberg University
Faculty of Chemistry and Geoscience
Institute of Geography

Master Thesis

Joint Use of Remote Sensing and Volunteered Geographic Information for Exposure Estimation

Author:
Anne Schauß
anneschauss@gmail.com

Supervisor:
Prof. Dr. Alexander Zipf
Dr. Hannes Taubenböck

May 2015

This master thesis was elaborated at the German Remote Sensing Data Center (DFD) of the German Aerospace Center (DLR), Oberpfaffenhofen
Department of Geo-Risk and Civil Security

Erklärung

Ich versichere, dass ich die Arbeit selbstständig verfasst habe, dass ich keine anderen Quellen und Hilfsmittel als die angegebenen benutzt und die Stellen der Arbeit, die anderen Werken dem Wortlaut oder dem Sinn nach entnommen sind, in jedem Fall als Entlehnung kenntlich gemacht habe. Das Gleiche gilt auch für beigegebene Zeichnungen, Kartenskizzen und Abbildungen.

Heidelberg, den

Abstract

The frequency and intensity of natural disasters have been increasing in the last decades due to raising vulnerability and an increasing amount of elements at risk with respect to hazardous conditions. Considering a high spatio-temporal variability of elements at risk, detailed information is costly both in terms of time and economic resources and therefore often incomplete and outdated. To alleviate these restrictions, the availability of very high resolution satellite images promotes accurate and detailed analysis over various scales, covering large geographical areas. To extract relevant information from the data, supervised classification and regression approaches are very popular due to their accuracy, robustness, and flexibility. The idea of such methods is to infer a rule (e.g., a decision function) from limited but properly encoded prior knowledge to assign e.g., a class label to unseen instances of the domain under analysis. Thereby, collection of appropriate prior knowledge (by e.g., detailed in-situ surveys) is normally the most time-consuming and expensive aspect with respect to data processing.

Meanwhile, the development of the Web 2.0 enabled Volunteered Geographic Information (VGI) to emerge as a new category of worldwide available geodata, which represents a new potential source of information. In this thesis, we develop a new methodology where VGI contributions to the OpenStreetMap (OSM) project are utilized as an alternative data source for gathering of appropriate prior knowledge. Based on a very high resolution image from the Worldview-2 sensor with 0.5 m spatial resolution, we provide a fusion scheme to automatically use the VGI information for a statistical learning approach. Thereby, we identify several land use land cover classes for the region of Valparaíso, Chile - an area highly prone to diverse natural hazards. Based on several regression techniques as well as disaggregation algorithms incorporating OSM data, additional information concerning exposed buildings and population is estimated.

The outcomes prove the applicability of VGI for remote sensing data processing. The exploitation of VGI is in particular important when proper prior knowledge is lacking or incomplete and could be imperative in the scope of an integrative risk assessment. The incorporation of VGI as source of essential prior knowledge disengages the common approach, which relies on site-specific dependencies, and promotes worldwide transferability and applicability.

Zusammenfassung

Die Häufigkeit und Intensität von Naturkatastrophen ist in den letzten Jahren aufgrund zunehmender Vulnerabilität und Anzahl gefährdeter Elemente deutlich angestiegen. Aufgrund der hohen raum-zeitlichen Variabilität gefährdeter Elemente ist die Beschaffung von detaillierten Informationen zeitaufwendig und kostspielig; vorhandene Daten sind deshalb oft unvollständig und nicht aktuell. Sehr hoch aufgelöste Satellitendaten reduzieren bisherige Restriktionen durch die Ermöglichung von präzisen, detaillierten Analysen über unterschiedliche Maßstäbe und großflächige Gebiete. Aufgrund ihrer hohen Genauigkeit, Robustheit und Flexibilität sind überwachte Klassifikations- und Regressionsverfahren zur Gewinnung von Informationen aus diesen Daten weit verbreitet. Die Grundidee solcher Methoden ist die Ableitung einer Regel (z.B. eine Entscheidungsfunktion) aus limitiertem, aber geeignetem Vorwissen, um auf diese Weise Instanzen ohne thematische Klassenkennzeichnung einer Klasse zuzuweisen. Bei der Datenprozessierung stellt die Akquisition von geeignetem Vorwissen den zeit- und kostenaufwendigsten Aspekt dar.
Währenddessen ermöglicht die Entwicklung des Web 2.0 nachhaltig die Entstehung von Volunteered Geographic Information (VGI) als eine neue Kategorie weltweit vorhandener Geodaten, welche eine neuartige potentielle Informationsquelle darstellen. In der vorliegenden Arbeit wird eine Methode entwickelt, in der VGI-Beiträge zum OpenStreetMap (OSM) Projekt als alternative Datenquelle zur Erfassung von adäquatem Vorwissen benutzt werden. Auf Basis eines sehr hoch aufgelösten Satellitenbildes des Worldview-2 Sensors mit 0.5 m räumlicher Auflösung wird ein Fusionskonzept zur automatischen Nutzung von VGI Informationen für einen statistischen Lernansatz bereitgestellt. Hierbei werden verschiedene Landnutzungs- und Bedeckungsklassen für die Region Valparaíso in Chile - ein gegenüber verschiedener Naturkatastrophen anfälliges Gebiet - identifiziert. Basierend auf unterschiedlichen Regressionstechniken sowie Disaggregations-Algorithmen werden durch die Integration von OSM-Daten zusätzliche Informationen bezüglich gefährdeter Gebäude und Bevölkerung abgeschätzt.
Die Ergebnisse zeigen das Potential von VGI für fernerkundliche Datenprozessierung. Die Nutzung und Integration von VGI gewinnt im Besonderen bei fehlendem und unvollständigem Vorwissen an Bedeutung, und könnte im Rahmen einer integrativen Risikoanalyse unerlässlich sein. Die Eingliederung von VGI als Informationsquelle für notwendiges Vorwissen löst den traditionellen Ansatz von standortspezifischen Abhängigkeiten und begünstigt somit weltweite Übertragbarkeit und Anwendbarkeit.

Acknowledgement

At this point I would like to thank all those who contributed to the success of this thesis.

First of all, I thank Dr. Prof. Alexander Zipf, who made this work possible as part of the Working Group Geoinformatics at the University of Heidelberg. Further, I particularly thank Jamal Jokar Arsanjani for his great cooperation in the advancement of this work and his valuable support and advices in the field of VGI and OSM within the course of the thesis.

Dr. Hannes Taubenböck for the friendly reception in his team "Modelling and Geostatistical Methods" at DLR and the facilitation of this thesis, as well as the good working atmosphere that he creates through his great motivation. Here I would like to especially thank Dr. Christian Geiß for the outstanding care and support. His valuable time, ideas, tips and discussions were indispensable for the success of this thesis.

Hendrik Zwenzner, without whom I would not have ended up at DLR (again). Thank you for supporting me since our first collaboration at DLR.

The DLR IT-support Fabian Henkel and Ralph Kiefl for their willingness to help with any technical problem.

Rudolf Richter for his great support regarding the atmospheric correction with ATCOR figuring out initially faulty satellite data.

Thomas Krauss for assistance in processing with CATENA.

Cristian Albornoz-Espinoza for his assistance in obtaining and processing the census data.

Ines Standfuß for the regular exchange of ideas and especially the collective maintenance of motivation in the basement.

Very special thanks are due to my parents for the unconditional and patient support in all aspects of life and especially for the motivation in the course of this thesis. And my brother for the support with any technical difficulty, proof reading and especially for his motivating distraction.

Contents

List of Figures

List of Tables

List of Abbreviations

AM	Area-Weighted Mean
ATCOR	Atmospheric and Topographic Correction
BL	Block Level
BUD	Built-up Density
CI	Contribution Index
CM	Confusion Matrix
CPP	Classification Postprocessing
DEM	Digital Elevevation Model
DI	Dispersion Index
DLR	German Aerospace Center
GHSL	Global Human Settlement Layer
GMESUA	Global Monitoring for Environment and Security Urban Atlas
GPR	Gaussian Process Regression
HOT	Humanitarian OpenStreetMap Team
IO	Image Object
IQR	Interquartile Range
JOSM	Java OpenStreetMap Editor
LC	Land Cover
LR	Linear Regression
LS	Largest Segment
LSM	Landscape Structure Metrics
LULC	Land Use and Land Cover
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
NB	Number of Buildings
nDSM	Normalized Digital Surface Model
NDVI	Normalized Differenced Vegetation Index
NIR	Near Infrared
NS	Number of Segments
OBIA	Object-based Image Analysis

OSM	OpenStreetMap
PA	Producer's Accuracy
PCA	Principal Component Analysis
POI	Point of Interest
RAE	Relative Absolute Error
RMSE	Root Mean Squared Error
RRSE	Root Relative Squared Error
RS	Remote Sensing
RTF	Rotation Forest
SAW	Simple Areal Weighting
SL	Segment Level
SMAPE	Symmetric mean absolute percentage error
SRTM	Shuttle Radar Topography Mission
SVM	Support Vector Machine
SVR	Support Vector Regression
UA	User's Accuracy
UNISDR	United Nations Office for Disaster Risk Reduction
VGI	Volunteered Geographic Information
VHR	Very High Resolution
WEKA	Waikoto Environment for Knowledge Analysis

1. Introduction

1.1. Background

The frequency and intensity of natural disasters have been increasing in the last decades due to raising vulnerability and an increasing amount of elements at risk with respect to hazardous conditions. Every day, millions of people all over the world are threatened by different natural (e.g., earthquakes, tsunamis, tornadoes) and man-made (industrial accidents) hazards [Pittore et al. 2015]. This illustrates the ongoing need for understanding assets at risk from such events [Ehrlich et al. 2012]. Causes for an increased risk are, for example, rapid urbanization and an increasing vulnerability of modern industrial societies. Economic, political, cultural and social pressures force people (especially in the Third World) to settle in high risk zones [Mueller et al. 2006].

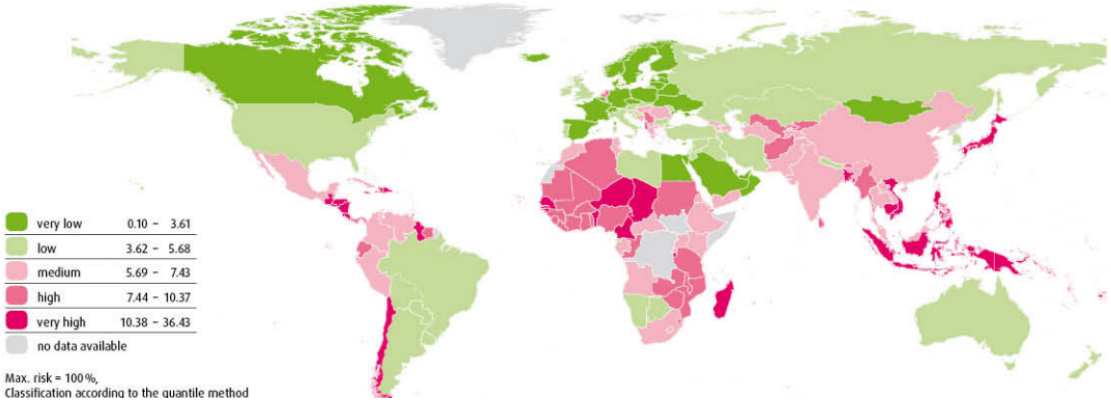


Figure 1.1.: WorldRiskIndex 2013: The WorldRiskIndex is calculated based on global available data and comprises 4 variables [Beck et al. 2012]: 1) Exposure towards natural hazards (e.g., eartquakes, cyclones), 2) Susceptibility depending on infrastructure, nutrition, housing situation and economic framework conditions, 3) Coping capacities depending on governance, disaster preparedness and early warning, medical services and social and material coverage, and 4) Adaptive capacities relating to forthcoming natural events, to climate change and to other challenges.

In this context, Chile ranks among the most by disasters affected countries with 61 reported events between 1980 and 2010, affecting more than 5 million people (Fig. 1.1) [Preventionweb 2015]. According to the World Risk Index, Chile further belongs to the 15 most exposed countries in the world and is assigned to the class of countries with the highest possible risk index (see Fig. 1.1). The index stresses, that not only the magnitude

of frequency of a natural event, but indeed also the social, economic and ecological factors characterizing a country essentially determine whether a natural hazard can turn into a disaster [Beck et al. 2012, Mucke et al. 2013].

While hazards can arise for many different reasons, the exposed assets (people, building,

No of events:	61
No of people killed:	1,581
Average killed per year:	51
No of people affected:	5,875,996
Average affected per year:	189,548
Economic Damage (US$ X 1,000):	33,049,160
Economic Damage per year (US$ X 1,000):	1,066,102

Table 1.1.: Natural disasters in Chile from 1980 - 2010 [Preventionweb 2015]

infrastructure) stay the same. For the mitigation, as well as the response of disasters, two basic data are essential: where the event is occurring and how many people and assets are likely to be affected [CEUDMT 2007]. The general distribution of potential hazard risk is often well-known. However, reliable, relevant and publicly available risk and exposure information is often missing and incomplete; and gathering appropriate information in terms of ground truth or reference datasets is expensive [Deichmann et al. 2011, Florczyk et al. 2014, Pittore et al. 2015]. This concerns in particular resource poor nations, that are at the same time often most vulnerable and exposed to the effects of natural or human induced events [Beck et al. 2012, CEUDMT 2007, Mucke et al. 2013]. Therefore, a crucial task in order to assess risk and exposure, is to identify and collect existing datasets that meet quality and semantic requirements [Florczyk et al. 2014].

The International Strategy for Disaster Risk reduction (ISDR) calls for a nationwide disaster risk assessment that should be precondition to disaster reduction measures [UNISDR 2015]. Therefore, public authorities are in particular need of vulnerability and exposure-related information in order to develop appropriate risk management and risk reduction strategies [Mueller et al. 2006]. Given the spatial variability in hazards, a need for spatially explicit exposure data is evident [CEUDMT 2007]. Thus, the main challenge is to find innovative, efficient methodologies to collect, organize, store and communicate exposure data, while accounting for the inherent spatio-temporal dynamics [Pittore et al. 2015]. Disaster management can only be as good as the available spatial information for decision makers [Taubenböck et al. 2007].

1.2. State of the Art

Information Extraction from VHR Imagery The identification of urban built-up areas is an important aspect in remote sensing-based applications and has gained attraction with the rise of very high resolution satellite images on the one hand [Herold et al. 2002a;b; 2003] and in terms of rapid urbanization on the other hand [Mueller et al. 2006, Shackelford and Davis 2003, Weng and Quattrochi 2008, Wurm et al. 2011]. Land use and land cover (LULC) mapping has been relying on data taken from many different sensors, but most of the recent efforts are related to very high resolution (VHR) satellite imagery, since it constitutes new potential for more detailed and accurate urban analysis [De Roeck et al. 2009, Herold et al. 2002b; 2003, Myint et al. 2011]. Particularly, the extraction of built-up masks has been one of the most interesting topics, since it forms the basis for further information retrieval such as exposure data [Liu et al. 2005, Wieland et al. 2012c]. During the past decade, considerable research has been done for automatic identification of urban areas from VHR satellite images that greatly facilitates the classification and extraction of buildings.

Thereby, spatial resolution is considered to be more important than spectral resolution [Herold et al. 2002b, Myint et al. 2011]. With VHR data, algorithms based on "per-object" segmentation rather than "per-pixel" classification are feasible [Blaschke 2003]. In this context, it is widely recognized that object-based image analysis (OBIA), using a two-step process, i.e. segmentation and classification, constitutes an appropriate approach [Chen et al. 2009, Giada et al. 2003, Herold et al. 2002b, Myint et al. 2011, Syed et al. 2005]. The aim of many recent algorithms is indeed to jointly consider area-based geometrical and spectral properties in order to recognize "objects" in the original image. Thereby, many different image segmentation algorithms exist and can be distinguished in the literature (e.g., edge detection and region growing) [Baatz and Schäpe 2000, Benediktsson et al. 2003, Blaschke 2003]. For the delineation of building footprints of varying scales, multi-scale optimization procedures are emphasized (hierarchical segmentation) [Ehrlich et al. 2012, Li et al. 2008, Peijun et al. 2007, Römer et al. 2012, Wieland et al. 2012c]. The incorporation of spatial, textural or morphological features in addition to spectral features for the classification of satellite images has been shown to overcome a lack of spectral resolution [Carleer and Wolff 2006, Giada et al. 2003, Li et al. 2014, Lari and Ebadi 2007, Singh et al. 2012, Tuia et al. 2009, Bellens et al. 2008]. Especially for urban scenes, geometrical and structural features are proposed to be useful rather than spectral information [Carleer and Wolff 2006, Ehrlich et al. 2012, Bellens et al. 2008, Lari and Ebadi 2007].

A different way to integrate contextual information is to extract shape information of single objects using mathematical morphology [Pierre Soille 2003]. Morphological oper-

ators have been recently used to classify panchromatic remotely sensed images and have been highlighted as very promising tools for data analysis [Benediktsson et al. 2003, Giada et al. 2003, Singh et al. 2012, Pesaresi and Benediktsson 2001, Soille and Pesaresi 2002, Tuia et al. 2009, Palmason et al. 2003]. Epifanio and Soille [2007] and Pina and Barata [2003] also use morphological analysis for image segmentation.
Several classifiers have been applied for interpreting landcover from VHR remote sensing imagery. Conventional parametric methods such as Bayesian classification, minimum distance, and maximum-likelihood proved to be simple and useful [Duda et al. 2001]. However, these approaches are based on rigorous statistical assumptions and require training data following a normal distribution. Therefore, non-parametric machine intelligence techniques have been increasingly incorporated into the classification of RS imagery, such as Support Vector Machines [Chi et al. 2008, Turker and Koc San 2010, Tuia and Camps-Valls 2009; 2011, Wieland et al. 2012c, Okujeni et al. 2013], Neural Networks [Rizvi and Mohan 2012, Lari and Ebadi 2007, Rizvi and Mohan 2010], and Decision Trees [Liu et al. 2005].
In general, supervised approaches have shown to be the most accurate and promising methodological solution for VHR imagery [Zhong and Zhang 2012, Paladini et al. 2011, Mianji and Zhang 2011, Samaniego et al. 2008, Del Frate et al. 2007, Giacco et al. 2010]. However, their applicability strongly depends on the availability of an exhaustive ground truth including samples from all the land cover classes present in the region of interest, which are often limited [Chi et al. 2008, Wieland et al. 2012c, Tuia and Camps-Valls 2011, Marconcini et al. 2014].

Exposure Estimation and Remote Sensing The increasing frequency and intensity of natural hazards [Manfré et al. 2012] raises the demand for exposure information as input variables for disaster risk models. Remote sensing is increasingly recognized as an important technique to derive exposure characteristics with respect to risk assessment [Ehrlich et al. 2012, Mueller et al. 2006, Eguchi et al. 2008, Geiß and Taubenböck 2013, Wieland et al. 2012c]. Thereby, high level of detail can be achieved by means of high resolution satellite imagery [Wieland et al. 2012b]. There are several studies that quantify extracted LULC classes in hazard prone areas, especially urban areas, for the derivation of exposure data [Mueller et al. 2006]. The development of automated image processing chains allows for full enumeration of the exposed assets over large areas while overcoming the time and cost constraints usually associated with manual workflows [Wieland et al. 2012c].
Ehrlich and Tenerelli [2013] address the use of remote sensing imagery to quantify the built environment and its spatio-temporal changes in order to analyze physical exposure. The extraction of building footprint maps, building location maps and built-up area maps

are identified for their use in the scope of disaster risk management. Aubrecht et al. [2013] emphasize the need of high-resolution census data that are independent of administrative units in order to quantify human exposure. Hereby, the population of interest is linked to the composition of the urban structure and the existing building stock [Wieland et al. 2012b].
In this context, Wieland et al. [2012c] estimate the number of buildings based on extracted built-up segments. They empirically derive a conversion factor per building type to count the actual number of buildings from the number of built-up segments, based on manually digitized reference-buildings. Ehrlich et al. [2012] use a combination of texture-based image processing, statistical sampling and photo-interpretation to quantify a building stock in terms of the number of buildings, the distribution of built-up areas and building size. Mueller et al. [2006] conduct fully- or semi-automatic inventory of buildings and their parameters in the scope of vulnerability assessment. Taubenböck et al. [2007] delineate an urban land cover classification based on IKONOS imagery to analyze the physical dimension of vulnerability (the spatial distribution of different types of buildings, the capacity of the street network or the identification of open spaces).
For the estimation of human exposure in terms of population distribution, many methods are reported in the literature [Wu et al. 2005]. Generally, studies on population estimation use available census data aggregated to large administrative units, to disaggregate and extrapolate population statistics based on remote sensing products (LULC classification) to more meaningful geographical entities [Wu et al. 2005, Taubenböck et al. 2009]. A widely used method in the scope of remote sensing is the dasymetric mapping approach, that uses ancillary data, such as population density in sub zones, landuse or building height information, as input [Freire and Aubrecht 2012, Wu et al. 2005, Mennis and Hultgren 2006, Steinnocher et al. 2011, Su et al. 2010, Wiesner and Taubenböck 2014].
For the purpose of estimating human exposure, Wieland et al. [2012c] for example disaggregate census data from higher administrative units to built-up units by estimating population density from the average floor area and the population per built-up unit. Further studies include also spatio-temporal population dynamics when assessing human exposure to hazards [Aubrecht et al. 2013, Freire and Aubrecht 2011; 2012, Freire and Gomes 2013]. There are various studies that extend remotely sensed derived exposure data by delineating further characteristics to describe vulnerability of these assets towards diverse natural hazards [Geiß et al. 2014a; 2015; 2014b, Mueller et al. 2006, Römer et al. 2012, Taubenböck et al. 2011; 2013, Wieland et al. 2012a]. For the case of Valparaíso, Indirli et al. [2010] conducted a risk assessment within the scope of the MARVASTO project, which is in particular concerned about cultural heritage protection [Indirli 2008]. They focus on the structural vulnerability of the historical center of Valparaíso and the

UNESCO protected churches located there.
These studies provide comprehensive analysis of exposure information, emphasizing the potentials of applying remote sensing techniques as a tool for supporting risk assessment at regional and local scales. However, there are still important challenges in this research field that need to be addressed: These include, among others, the development of transferable methods on the one hand and to improve the accuracy of results derived from remote sensing on the other hand. Data availability is the main limiting factor for both issues [Römer et al. 2012, Wieland et al. 2012c].
The integration of user-generated geographical content, termed as Volunteered Geographic Information (VGI) [Goodchild 2007a] is recently suggested as a potential opportunity and complementary source of information for classification purposes as well as exposure assessment [Goodchild 2007a, Goodchild and Glennon 2010, Kinley 2013, Pittore et al. 2015]. The added value VGI can bring to remote sensing analysis is increasingly recognized and accepted in the RS community in recent years [Arsanjani et al. 2013, Pittore et al. 2015, Kinley 2013]. Yet, little work exists combining traditional and crowd-sourced data for information retrieval or exposure estimation.

Volunteered Geographic Information and Remote Sensing The availability of spatial data on the web provided by crowdsourcing communities, as well as its scientific acceptance has increased significantly in recent years. The integration of VGI in data analysis provides promising opportunities due to the amount of information made available [Goodchild 2007a, Kinley 2013, Sester et al. 2014]. Yet, it also poses new challenges in terms of heterogeneity of the data and quality assurance [Flanagin and Metzger 2008, Haklay 2010]. Therefore, the focus in recent research lies at discerning the quality of VGI and hence to identify its applicability to further projects by complementing or substituting commercial data [Arsanjani et al. 2015, Ather 2009, Ciepłuch et al. 2011, Fan et al. 2014, Haklay 2010, Hecht et al. 2013, Helbich et al. 2012, Kounadi 2009, Mooney et al. 2010, Neis and Zipf 2012, Zielstra and Zipf 2010]. Yet, the integration of non-authoritative data with traditional, well established data and methods is a novel approach [Kinley 2013, Schnebele and Cervone 2013]. Very little research has been conducted in terms of combining VGI with remote sensing techniques. Hence VGI has rarely been incorporated within classification algorithms used to generate thematic maps [Kinley 2013]. However, some recent studies take verification a step further and use VGI as an additional data source to refine or update existing information.
Schnebele and Cervone [2013] fuse remote sensing and VGI data in order to introduce a new methodology for the generation of flood hazard maps. They use different layers generated based on Landsat ETM+ images, DEM and finally VGI to refine the information. Here, VGI data include especially videos and photos from Flickr, YouTube, Weather Un-

derground, Wikipedia and abc24.com. Their methodology is particular useful if satellite data is limited or of poor quality. Klonner et al. [2014] fuse OSM with Airborne Laser Scanning (ALS) data. OSM polygons are taken as a mask to extract nDSM values of existent ALS buildings and a new fixed building height is integrated into the nDSM of existent ALS buildings where no building is present. Kunze and Hecht [2015] develop a method to enrich official building footprints with semantic VGI information from the OSM project in order to refine the estimation of the number of dwelling units and population using building-based dasymetric mapping techniques. In these studies, existing information are updated and enriched by VGI rather than substituted.
In contrast, Arsanjani et al. [2013] implement VGI contributions to the OSM project as alternative source of information for training samples instead of field measurements. Very high resolution RapidEye satellite imagery is used. For this purpose, they harmonize OSM features with the GMES Urban Atlas classes to be able to validate the methodology. Furthermore, training sites are checked for outliers that are erased from the database. With a κ statistics of 89% they prove that VGI is an appropriate data source for the introduction of training sites to the remote sensing process.
Moreover, there are several studies using VGI and in particular OSM for validation purposes [Foody and Boyd 2013, Pesaresi et al. 2013]. The probably biggest project is the Global Human Settlement Layer (GHSL), a satellite imagery processing framework for the derivation of a map of European settlements [Pesaresi et al. 2013, Florczyk et al. 2014]. OSM derived reference data is used within the framework at several stages of an information layer production, namely learning, classification, post-processing and validation [Florczyk et al. 2014].
Throughout several publications, the potential and increasing acceptance of VGI in science is emphasized [Kinley 2013, Sester et al. 2014]. In particular in the domains of disaster and community mapping, VGI has proven and proves to be useful [Goodchild and Glennon 2010, Pittore et al. 2015, Poser and Dransch 2010]. In summary, VGI has primarily been used in terms of additional datasets augmenting information content rather than substituting essential data input. Arsanjani et al. [2013] published the first trials to combine and integrate traditional and crowd-sourced data in a remotely-sensed classification approach. Their achieved results pose the feasibility of the incorporation of VGI in remote sensing and emphasize the potential of further research.

1.3. Research Objectives

The increasing demand for exposure information requires to effectively map the built-up environment and its changes for regions that experience rapid urbanization or are located in hazard prone areas [Ehrlich et al. 2012]. For classification issues, supervised learn-

ing techniques generally constitute the most accurate methodological solution in remote sensing [Zhong and Zhang 2012, Paladini et al. 2011, Mianji and Zhang 2011, Samaniego et al. 2008, Del Frate et al. 2007, Giacco et al. 2010], however, their applicability strongly depends on the availability of an exhaustive ground truth, which is seldom available because of its high costs [Marconcini et al. 2014]. In particular in the case of big data scenarios (e.g. high resolution data processing) costs rise to unacceptable heights and therefore data availability features the limiting factor for an integrative risk assessment. Hence, the goal is to find new possibilities in order to gather in-situ information and data in conjunction with low expense. In the Global Assessment Report on Disaster Risk Reduction 2015 (GAR), VGI is proposed as a potential opportunity and complementary source of information for exposure assessment for the first time [Pittore et al. 2015]. Indeed it is a powerful new development providing a great amount of up-to-date and area-wide geographic information. The exploitation of the potential of this exhaustive data source for remote sensing applications constitutes an extraordinary opportunity and urgently needs to be investigated.
Therefore, the aim of this study is to exploit the potential of VGI in remote sensing data processing. A methodology will be developed where VGI contributions to the OpenStreetMap (OSM) project are utilized as an alternative data source for gathering appropriate prior knowledge in the scope of supervised landcover classification and exposure estimation. Thereby the following research questions will be tackled:

- How good can LULC patterns be mapped by deploying OSM contributions as alternative information source for gathering training samples instead of traditional in-situ data collection?
- What is the potential of OSM to derive further exposure estimation?

By means of a supervised approach based on worldwide available ground-truth data, transferability will be facilitated and applicability made feasible for diverse other thematic and geographical areas. In terms of the integration of VGI within a classification framework, OSM data are intended to be utilized to generate reliable training data at low cost and without in-site visits, tackling the problem of low availability of ground-truth. An important step will be the quality and plausibility assessment to ensure reliability of training input since it has a great impact on the feasibility and outcome of the analysis. The added value of this study will therefore be the exploitation of exhaustive, until now unavailable data sources for remote sensing applications. The generation of training data based on OSM, and especially the data fusion and/or methodological integration, is central. In this context, it is aimed at integrating VGI not only as an additional source of information, but to rather entirely substitute and facilitate the expensive and time-consuming process of determining the essential ground truth data.

Outline

This work is organized as follows. The thesis starts with a basic introduction of the fundamental terminologies concerning exposure, remote sensing image analysis and VGI in chapter 2. Chapter 3 presents the study site and provides an overview of the underlying spatial data. Chapter 4 comprises all implemented working steps in this study. Starting with relevant pre-processing steps it is lead to the proposed approaches for information extraction by supervised object-based classification, incorporating OSM-data as training samples. Subsequently, the focus is shifted to exposure estimation and methodologies illustrated for both, the estimation of number of buildings and exposed population. In chapter 5 results are presented and discussed, a conclusion and outlook are given in chapter 6.

2. Theoretical Background

The following chapter gives an overview and a basic introduction to the fundamental terminologies concerning exposure, remote sensing image analysis and Volunteered Geographic Information (VGI).

2.1. Exposure to Natural Hazards

Conceptual framework of hazards, vulnerability and risk In terms of the conceptualization of risk, a shift can be observed within the scientific community from hazard-oriented strategies towards more integrative approaches. This is due to the realization that natural hazards do not have an intrinsic dangerous character themselves [Turner et al. 2003, Wisner et al. 2003, White et al. 2005]. Nowadays, the prevalent definition of risk can be described on the basis of two factors. According to UNISDR [2004], risk is a function of hazard and vulnerability (Eq. 2.1). A hazard constitutes a potentially damaging event whose magnitude is dependent from the likelihood of occurrence, intensity, frequency and location. In contrast, vulnerability describes the degree of susceptibility of the elements exposed to a particular event. In this context, risk is described as a relationship between the degree of exposure and the degree of damage [UNISDR 2004].

$$\text{Risk} = \text{Hazard} * \text{Vulnerability} \tag{2.1}$$

This definition can be extended by describing vulnerability in its further subcomponents [Taubenböck 2008, White et al. 2005]. Equation (2.2) defines vulnerability as the interrelation of the exposure and the susceptibility as stressor of the system, and the capacity of coping with a stressor as mitigation factor [Taubenböck 2008].

$$\text{Vulnerability} = \frac{\text{Exposure} * \text{Susceptibility}}{\text{Coping Capacity}} \tag{2.2}$$

According to Bohle [2001], vulnerability accounts for a double structure with exposure as the external and coping capacity as the internal side. Table 2.1 shows the dimensions and their manifestation contributing to a holistic perspective of vulnerability in the risk framework.

RISK			
HAZARD		**VULNERABILITY**	
	Exposure	**Susceptibility**	**Coping Capacity**
Earthquake, Tsunami, Landslide, etc.	degree of exposed built-environment	height, density, number of buidlings, age etc.	building codes, protection measures
	location	accessibility, distances, height, slope	urban planning
	% of population living in high risk zone	total population, population density, gender, etc.	evacuation plans, access to information, etc.

Table 2.1.: Concept of risk, hazards and vulnerability [Taubenböck et al. 2007]: Risk is a function of the hazard and the given vulnerability. Vulnerability is composed of the subcomponents Exposure, Susceptibility and Coping Capacity whose relationship is depicted in equation (2.2).

Exposure can be defined as the degree, duration and/or extent in which a system is in contact with, or subject to, perturbation. According to Pittore et al. [2015] exposure comprises "all elements that may be affected by one or more hazards including lifelines (transportation, power, water, etc.), buildings, population, societal functions, the economic system, and cultural heritage". Exposure estimation is an essential part of an integrative risk assessment and is usually considered the starting point [Ehrlich et al. 2012]. Within this context, measures of exposure can include a quantitative assessment of types of assets or the number of people in an area at risk [Wieland et al. 2012b].

Thus, the central objective of risk assessment is to provide indications where, when and how people and associated assets might be affected by a certain impact [Aubrecht et al. 2013]. Its result should provide decision-and policymakers with supporting information to target response and mitigation actions adequately [Aubrecht et al. 2013]. Depending on the scale of analysis, exposure data may vary from detailed descriptions of characteristics and location of structures that may be damaged (e.g. buildings), up to aggregated composite models for larger geographical entities (e.g. cities).
The location and characteristics of exposed assets do not only provide a crucial basis for risk modelling, but also support damage and economic loss calculations from extreme natural events [Cutter 1996]. When exposure information is estimated, two strategies can be differentiated: Statistical sampling and full enumeration [Pittore et al. 2015]. Sampling analyzes only small subset areas to subsequently estimate summary statistics of the whole city. In contrast, full enumeration from satellite imagery refers to the detection and definition of each exposed asset within the study area, and can achieve high levels of detail [Pittore et al. 2015, Wieland et al. 2012b].

2.2. Remote Sensing Data Analysis

Very High Resolution RS Imagery In recent years, the availability of very high resolution space images has been increasing fast. The first successful launch of commercially developed high resolution earth observation satellites occurred in 1999 with the launch of Space Imaging's IKONOS system providing a 82-cm panchromatic and 4 m multispectral resolution [Lillesand et al. 2004].
High spatial resolution satellite sensors have substantially increased the potential for detailed digital imaging of complex structures, especially urban applications [Herold et al. 2002b, De Roeck et al. 2009]. Yet, despite the advantage of the higher level of spatial detail offered by new sensors, they pose new challenges and requirements for the analysis at the same time [De Lange 2013, De Roeck et al. 2009]. The internal spectral variability (intraclass variability) of each landcover class increases and decreases the spectral variability between different classes (interclass variability) [Bruzzone and Carlin 2006]. The high intraclass and low interclass variabilities lead to a reduction of the statistical separability of the different landcover classes in the spectral domain, which in turn involves high classification errors [Bruzzone and Carlin 2006].
Urban areas are especially complex spatial environments, typically consisting of built-up structures, various vegetation covers, bare soil zones and water bodies [Herold et al. 2002b]. Different objects in urban areas (e.g., buildings) do not present homogenous characteristics, which signifies that the same class can show very different properties due to different disturbances. One of these disturbances is the relief of the terrain and skewed camera angles which lead to unequal illumination of the surface and therefore inhomogeneities in the multispectral data [Albertz 2013]. Due to the high spatial resolution of the images, pixels are mostly smaller than the urban objects that need to be distinguished [De Roeck et al. 2009]. The detection of buildings for example becomes a challenging task due to complexity and similarity between buildings and other similar areas like roads, cars or bare soils. Hence, context-based approaches that take into account not only spectral but also spatial characteristics of the scene contribute to a better separation of spectrally similar surface types, and therefore constitute the appropriate approach for VHR Imagery [De Lange 2013, De Roeck et al. 2009].

Object-based Image Analysis (OBIA) One of the most promising approaches in order to counter the challenge of efficiently extracting information from continuously improving spatial resolution RS imagery is the so-called object-based image analysis [Blaschke 2003]. OBIA uses a two step process instead of the one step process of the traditional pixel based image classification techniques [Campbell and Wynne 2011]. Primarily, image objects are generated based on homogeneity criteria resulting in objects of similar char-

acteristics. There exist various algorithms and techniques to derive homogenous image objects [Blaschke 2003]. In this study, the process is referred to as segmentation.
These homogenous regions form, rather than individual pixels, the entities to be classified in the second step. As the objects are composed of many pixels, they have a multitude of implicit and derivable properties, which in this study will be referred to as features [Campbell and Wynne 2011].
Features can describe the geometry, texture, spread, central tendency, neighborhood relations or any other contextual information for instance [Ehrlich et al. 2012]. In this context, mathematical morphology is a powerful image analysis technique to analyze shape and form of objects [Pierre Soille 2003]. Morphological image filtering is useful to reduce the complexity of the image, to extract or suppress image objects or structures [Acar and Bayram 2009, Pierre Soille 2003]. In the classification process the numerous features which characterize each, during the segmentation process derived image object, can be used for the distinction of different land cover classes [Campbell and Wynne 2011].

Classification of RS Imagery In remote sensing, numerous classification techniques exist for the purpose of pattern recognition. Machine learning is an area of artificial intelligence and generally refers to the development of methods that optimize their performance iteratively by learning from data [Waske et al. 2009]. Machine learning techniques can be predictive or descriptive (Fig. 2.1). Predictive models can make predictions of a specific phenomenon, e.g. predicting population densities by regression, and descriptive models generally aim at distinguishing classes of different patterns, e.g. landcover classifications [Waske et al. 2009].
Additionally, it can be distinguished between supervised and unsupervised methods, depending on whether a priori knowledge (labeled training samples) during the decision process is available or not [Waske et al. 2009]. Supervised approaches constitute the most accurate methodological solution due to the supply of ancillary information, however, their applicability strongly depends on the availability of an exhaustive ground truth [Marconcini et al. 2014]. Supervised learning can be divided into regression and classification issues. Thereby, the outputs for classification are discrete class labels while regression concerns the prediction of continuous quantities [Rasmussen and Williams 2006]. Supervised algorithms learn from a set of training samples whose class membership is known in order to assign the target data samples to a class. Here, a trade-off is made between how well the function fits the data and how complicate the function is [Waske et al. 2009].

Furthermore, it can be differentiated between parametric and non-parametric approaches. Parametric methods are based on statistical parameters (mean & standard deviation) assuming that the data follow a specific distribution. In contrast, non-parametric approaches are not constrained to prior assumptions on the distribution of the input data.

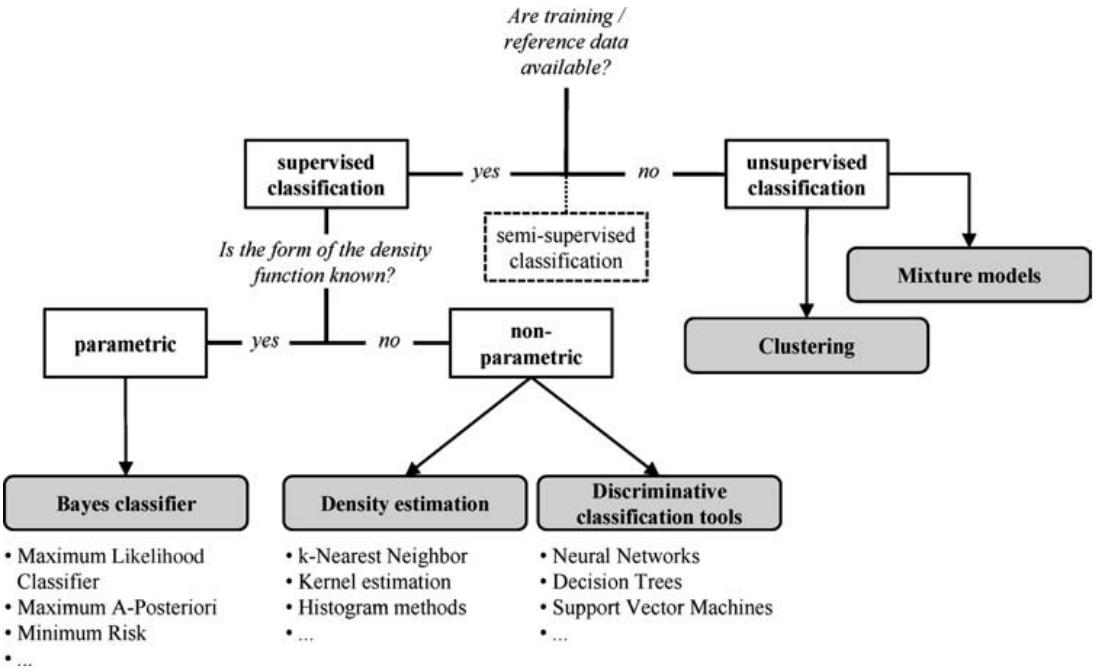


Figure 2.1.: Categories of classifiers [Waske et al. 2009]: Supervised and unsupervised classification approaches can be distinguished by whether labeled training instances are available or not. Parametric approaches assume the data to follow a specific statistical distribution while non-parametric methods are not constrained to prior assumptions on the data distribution and are rather based upon objects in a feature space. Supervised, parametric approaches comprise the Bayes classifiers while non-parametric approaches can be further separated in density estimations and discriminative classification tools.

They are rather based on objects in a feature space. However, for more complex data, this approach is more appropriate since the class distribution can often not be modeled by adequate multivariate statistical models; additional information like shape descriptors will be helpful.

Decision trees refer to a hierarchical model of numerous splits, each consisting of a binary decision and its consequences. The input data is passed through the branches, ending in a leaf, which represents one of the target classes. The dataset is recursively partitioned into smaller subdivisions, separating the data that meets the formulated condition from the rest of the data [De Roeck et al. 2009]. Decision trees can be used for both, classification (classification tree) and regression tasks (regression tree) [Waske et al. 2009, Rokach and Maimon 2008]. Their strength lies in their flexibility, non-parametric nature and ability to handle nonlinear relations between features and classes [Kavzoglu and Colkesen 2013].

Ensemble learning is a machine learning paradigm where multiple learners are trained to solve the same problem. In contrast to ordinary machine learning approaches which try to learn one hypothesis from training data, ensemble methods try to construct a set of hypotheses and combine them [Zhou 2009]. Classifier ensembles or multiple classifier

systems have been used successfully in recent years, in particular for multi-source and high-dimensional data [Waske et al. 2009]. The idea behind is combining different classifier methods or variants of the same classifier and determine results by a simple voting scheme, and thereby, improving the classifier performances. It is assumed that each independent classifier produces errors which are not produced by the majority of the other classifiers, and therefore, can be eliminated by the combination of the individual outputs. One of the recently introduced ensembles is the rotation forest which has been applied in this study. It is based on the idea of building accurate and diverse classifiers by applying feature extraction on the training sets and then reconstructing new training sets for each classifier [Kavzoglu and Colkesen 2013].

2.3. Volunteered Geographic Information

During the past few years, the rise of the Web 2.0 in conjunction with increasing numbers of internet users worldwide have enabled a rapid increase of the amount of social networks and user-generated content in the internet [Kamel Boulos et al. 2011]. Web platforms and Wikis are not only used to gather information but also to add data and edit the provided content of others [Goodchild 2007a]. Within Geographical Information Science (GIScience) research, this type of voluntary information is referred to as "Volunteered Geographic Information" (VGI) [Goodchild 2007a]. It combines the terms Web 2.0, collective intelligence and neogeography and can therefore be seen as a special case of the user-generated content in the Web [Goodchild 2007b].
Volunteered Geographic Information (VGI) is about the widespread engagement of a large number of citizens, mostly without specific formal qualifications, creating, collecting and sharing geographic information voluntarily [Goodchild 2007a]. VGI is obviously enabled by technologies like Web 2.0, broadband communication, location-aware mobile devices, GPS, etc. [Goodchild 2007a]. In contrast to authoritative data, VGI do not rely on a standardized vocabulary and are not completely reliable. The data acquisition depends rather on the availability of information and volunteers willing to contribute. Therefore, the quality results in a substantial inhomogeneity. Quality assurance is performed by the users themselves and not by a detailed quality assurance scheme ensuring that the data conforms to given specifications like it is the case for authoritative data [Sester et al. 2014]. VGI poses new opportunities for GIScience and cartography, but raises also new challenges, especially with respect to data quality and redundancy [Sester et al. 2014].

The OpenStreetMap Project (OSM) is one of the most popular and well supported examples of VGI [Mooney et al. 2010]. The project was launched in London in August 2004 with the aim to create a free digital editable vector map of the world [Haklay 2010].

Information is collected by many participants, collated on a central database and distributed in multiple digital formats through the World Wide Web as "Open data" according to the Open Data Commons Open Database License (ODbL) [OSM 2015c]. Over the last years, the number of registered users in OSM has grown continually [Neis 2015]. From 50,000 at the beginning of 2012, the number of registered users raised up to more than 1,900,000 by 2015 (Fig. 2.2). The contributors are not necessarily experts, in most cases

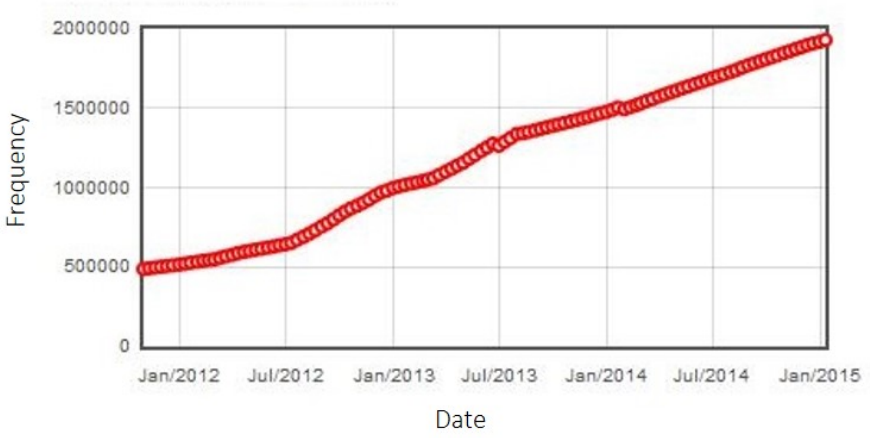


Figure 2.2.: Number of registered OSM users. The statistics were retrieved from the free wiki world map (OSM) and processed by Altogetherlost [2014] and Neis [2015].

they are amateurs [Haklay et al. 2010]. The standard means of collecting and uploading data to OSM is by collecting data using GPS devices or digitizing of locations from publicly available aerial imagery [Mooney et al. 2010]. For this purpose several companies support the project (e.g. Bing, Yahoo) [Neis and Zipf 2012].
There are no compelling rules of how to edit OSM-features, but contributors are encouraged to use the vocabulary suggested in the OSM-Wiki. Thus, the quantity and quality of OSM data is very heterogeneous and depends mainly on the contributors accuracy and experience, as well as on the number of active contributors in the area of interest [Haklay 2010]. The majority of members is located in Europe (72%) and European countries therefore belong to the most developed countries in terms of completeness of provided OSM data [Neis and Zipf 2012]. In urban areas OSM-data competes with commercial data due to great detail, and often even higher degree of detail and information density [Zielstra and Zipf 2010, Neis and Zipf 2012, Hecht et al. 2013]. In contrast, rural areas and especially developing countries feature incomplete and often very inaccurate data provision.

3. Study Site and Spatial Dataset

The following chapter outlines the study area Valparaíso in Chile with its characteristics, especially regarding the townscape, city structure and proneness to natural hazards. Its geographical and geological setting exposes Valparaíso to various natural hazards and make it a relevant study area in terms of an integrative risk assessment. In addition, section 3.2 introduces and describes the utilized datasets that constitute the base for investigations throughout this study.

3.1. Valparaíso, Chile

As study site the city of Valparaíso in Chile was chosen due to its high susceptibility to diverse natural hazards. Valparaíso, the second largest city in Chile, is located 120 km northwest of the capital Santiago on the coast at a latitude of 33°02'21" S and a longitude of 71°37'38" W (Fig. 3.1) . It is characterized by a mediterranean climate with dry summers, heavy rainfalls in winter and fog due to the Humboldt Current throughout the whole year [Erikson and Högstedt 2004]. In the 2002 census 275,982 inhabitants were counted, however, unpublished numbers of the controversial 2012 census indicate 284,630 inhabitants [Munivalpo 2015, INE 2015]. The city has been steadily growing, whereby the growth rate is decreasing in recent years as figure 3.2 suggests. Because of its geographic position, growth is limited by the sea and rather uninhabitable hilly regions.

The city is highly prone to various natural hazards (seismic events, tsunamis, landslides, etc.) and anthropic calamities (mainly wild and human-induced fires), which make Valparaíso a paradigmatic study case about hazard mitigation (Fig. 3.1, c) [Indirli et al. 2010; 2008, Indirli 2008]. Whole Chile is threatened by seismic and tsunami hazard due to the subduction of the three major plates (Nazca Plate, South American Plate, Antarctic Plate) at the west border of the country. However, the geographical (located right at the sea) and geological (seismic amplification due to soft soil) setting makes Valparaíso especially prone to them. Since the first systematic seismological recorded earthquake in Chile that was also the most disruptive one (Magnitude Richter 8.3) devastating Valparaíso in 1906, there were 12 earthquakes affecting Valparaíso with intensities higher than 7 reported until 1985 [Indirli et al. 2010]. The majority had their epicenter offshore of Valparaíso region along the plate boundary. In three of these cases there was also a

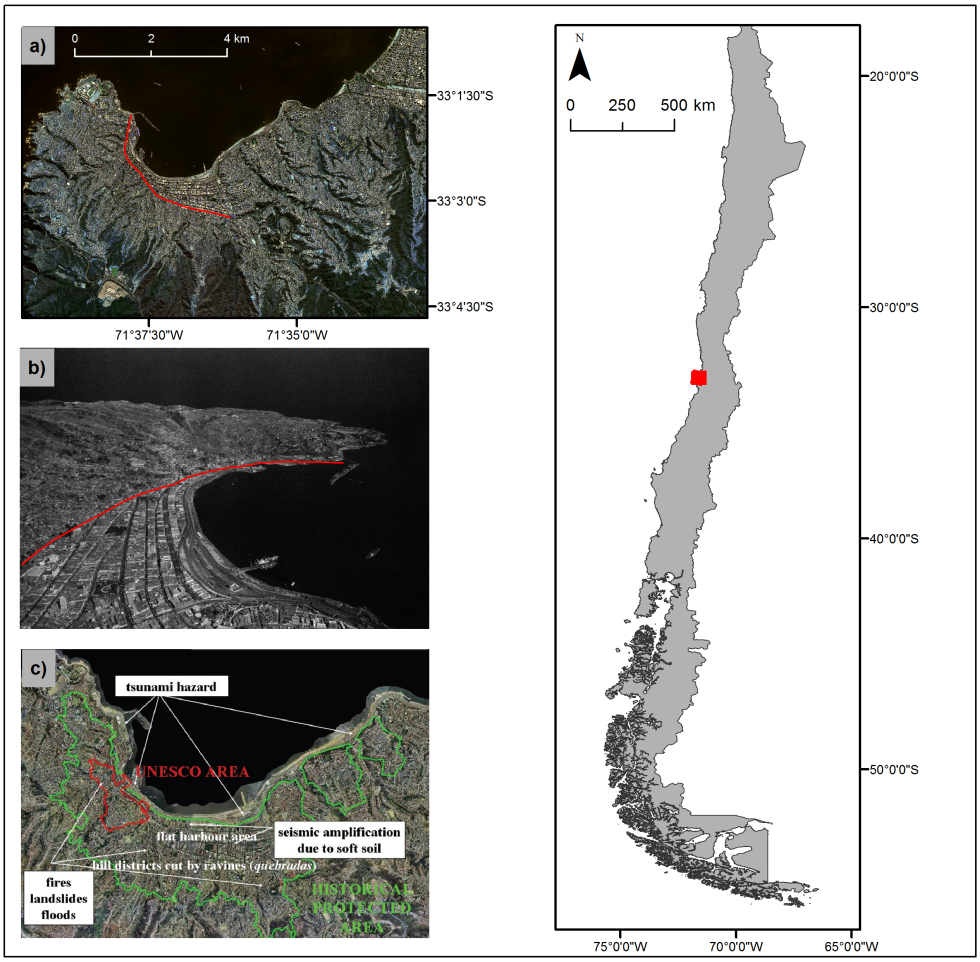


Figure 3.1.: Studysite Valparaíso. a) Worldview-2 scene Valparaíso, red line: seperation flat harbour areas and hill quarters b) Components of topography [Sánchez M. et al. 2009], red line: seperation flat harbour areas and hill quarters c) Hazards and safeguarded areas, green area: historical protected area, red area: UNESCO protected historical center [Indirli et al. 2008]

tsunami generated (1906, 1918, 1985). The earthquakes often resulted in devastation, injured, homeless people and dead [SHOA 2015].

Furthermore, there is a high and dangerous landslide risk especially in case of great and fast rainfalls due to soil material (schist basement, colluvial layers, fluvial terraces) and the exposure of slopes as well as existing deep fluvial incisions ("quebradas") [Erikson and Högstedt 2004, Indirli et al. 2008]. Several landslides affecting houses or roads are reported every year [Erikson and Högstedt 2004]. Fire risk in the area is worsened by usual windy weather, the narrow and tortuous hill roads, presence of mainly wooden houses and sometimes insufficient water pressure in the hydrants [Indirli 2008]. A recent example are the forest fires that occurred in April 2014 which could not be stopped for several days (April 12th - 17th). Almost 1,000 hectares were burnt, devastating about

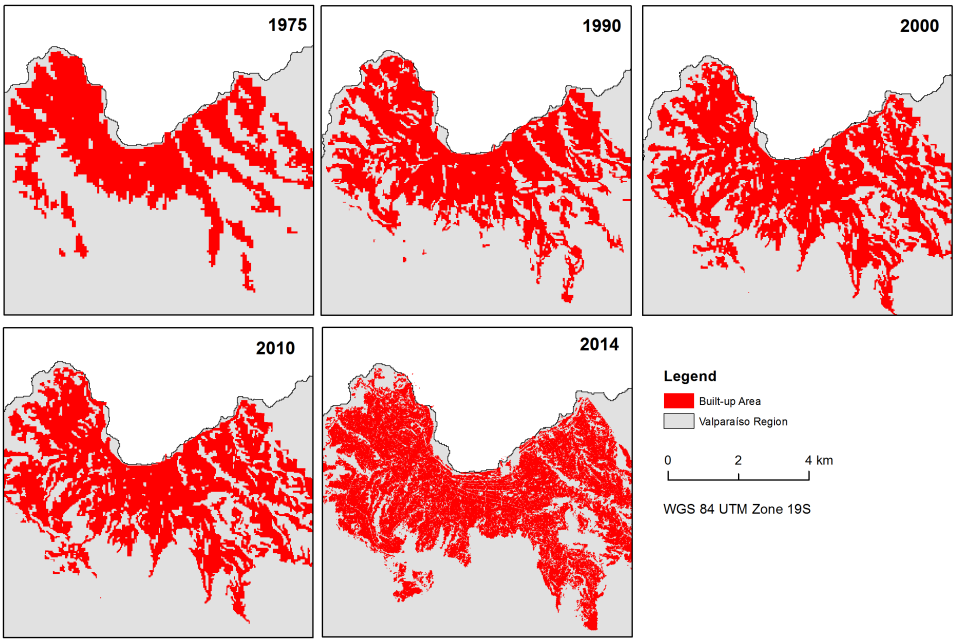


Figure 3.2.: Expansion of Valparaíso from 1975 - 2014: Built-up masks derived from LANDSAT MSS, 60 m resolution (1975), LANDSAT TM, 30 m resolution (1990, 2000), LANDSAT ETM+, 30 m resolution (2010) and Worldview-2, 0.5 m resolution resampled to LANDSAT 30 m resolution (2014)

3,000 houses in 12 districts ("barrios"), affecting 13,000 people [ONEMI 2014].
The important morphology of Valparaíso with respect to the risk factors can therefore be roughly divided in two characteristic sectors (Fig. 3.1 a,b):
Firstly, the flat harbour area, characterized by port facilities up to the waterfront and the commercial district including rather big buildings, straight streets, highways and rail tracks parallel to the coast. Secondly, the hill quarters, covering 49 hills, cut by valleys, and climbed by narrow and snaky lanes, filled with small and squat houses, typically made of wooden frames, adobe panels and covered by zinc tinplate ("calamine") [Indirli et al. 2008]. The historical center which is located between the sea and the first terrace is included since 2003 in the UNESCO World Heritage List of protected sites (see Fig. 3.1,c) [UN 2003].

3.2. Spatial Dataset

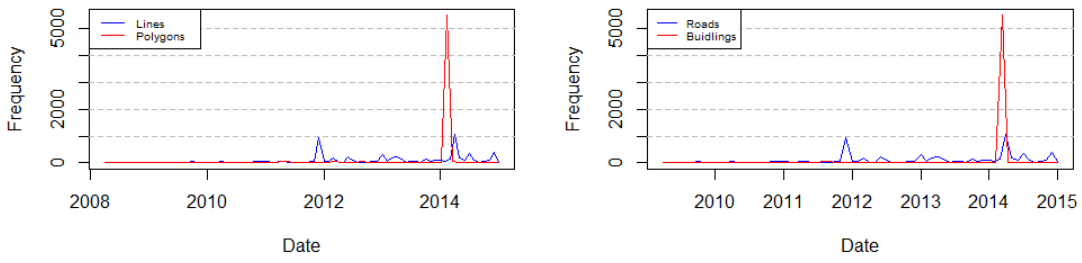
3.2.1. Worldview-2

Very high-resolution satellite imagery from the Worldview-2 sensor, launched in October 2009 and supplied by Digital Globe, Inc, is provided [Digital Globe 2015]. The system

operates on a sun-synchronous orbit at an altitude of 770 m. The average revisit time is 1.1 days, depending on latitude and image collection angle. Worldview-2 provides a 0.46 m panchromatic resolution and a 8-band multispectral sensor with the resolution of 1.85 m [SIC 2015]. The 4-band bundle is available for this thesis, aquired on April 19, 2014 at 15:01 p.m. The spectral resolution of the 4 bands are: Blue: 0.450 - 0.510 μm, Green: 0.510 - 0.580 μm, Red: 0.630 - 0.690 μm and NIR: 0.770 - 0.895 μm.

3.2.2. OpenStreetMap

For automated generation of training samples the OpenStreetMap shapefiles for Chile from November 4, 2014 were aquired from the official data provider Geofabrik [Geofabrik 2015]. The data are structured as Points of Interest (POI), places, roads, railways, waterways, natural, land use, and buildings [OSM 2015d]. Information about the history of the objects can be extracted from the OSM-Planetfile [OSM 2015e]. The OSM-Planetfile stores all nodes (configuring points, polylines and polygons) ever contributed to OSM. The additional attributes "osmuser" (User name), "osmuid" (User ID), "osmversion" (determines how many times a node has been edited) and "osmtimesta" (date node has last been edited) give information about the history of every object. The analysis of the



(a) Lines and polygons

(b) Roads and buildings

Figure 3.3.: Edited OSM-Nodes in the study area until January 2015: Almost all added polygons and lines are roads and buildings.

history reveals that most objects within the study site were added in April 2014 whereas before and after contributions were quite low. In all other months, apart from April 2014 the monthly average of edited lines and polygons were 78 and 36 respectively, whereas in April 1,085 lines and 13,339 polygons were added. All added line features are roads and 98% of added polygons are buildings (Fig. 3.3). This suggests that only 2% of the polygons comprise landuse information other than buildings. The amount of editing users

per month also shows a great peak in April 2014 (Fig. 3.4).

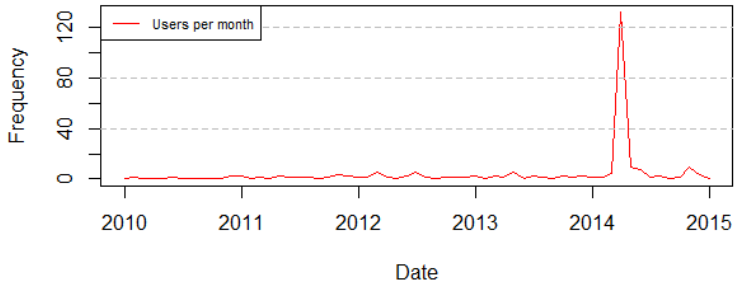


Figure 3.4.: Number of editing users per month from 2010-2015. A peak is recorded in April 2014 while throughout all other months the number of active users does not exceed ten users per month.

The time frame when most editing was done is exactly at the time when Valparaíso was hit by devastating fires affecting several neighborhoods between 12th and 17th April 2014 (Sec. 3.1). Further analysis shows, that the majority of the houses for the study area were generated within the scope of two HOT-OSM Tasks with the aim to "rapidly map the neighborhoods affected by the disaster as they were before the fire, so that this data could be used to assess damages and assist with humanitarian response activities" [OSM Task Manager 2014a;b]. The task instructions request the volunteers to use Bing imagery for mapping [OSM Task Manager 2014a;b, Global Resources News 2014].
The Humanitarian OpenStreetMap Team (HOT) acts as a bridge between the traditional Humanitarian Responders and the OpenStreetMap Community to provide up-to-date remote data processing during disasters [HOT OSM 2015].

3.2.3. Census Data

Census data for the city of Valparaíso are obtained from the National Statistics Institute of Chile (INE). Population statistics are available in tabular form and census units independently as shapefiles. Starting in 1813 approximately every 10 years a census was conducted in Chile. The last census was carried out in 2012, but the data has been retracted due to a series of problems during acquisition and processing [Bianchini et al. 2013]. This emphasizes the general problem of trustworthiness of official census data. The only official released data for 2012 are based on municipality ("Comuna").
For 2002, data is available in all resolutions following administrative and electoral division boundaries. These are, ordered by decreasing size, "Comuna" (municipality), "Distritos censales" (census districts), "Zonas censales" (census zones) and "Manzanas" (building

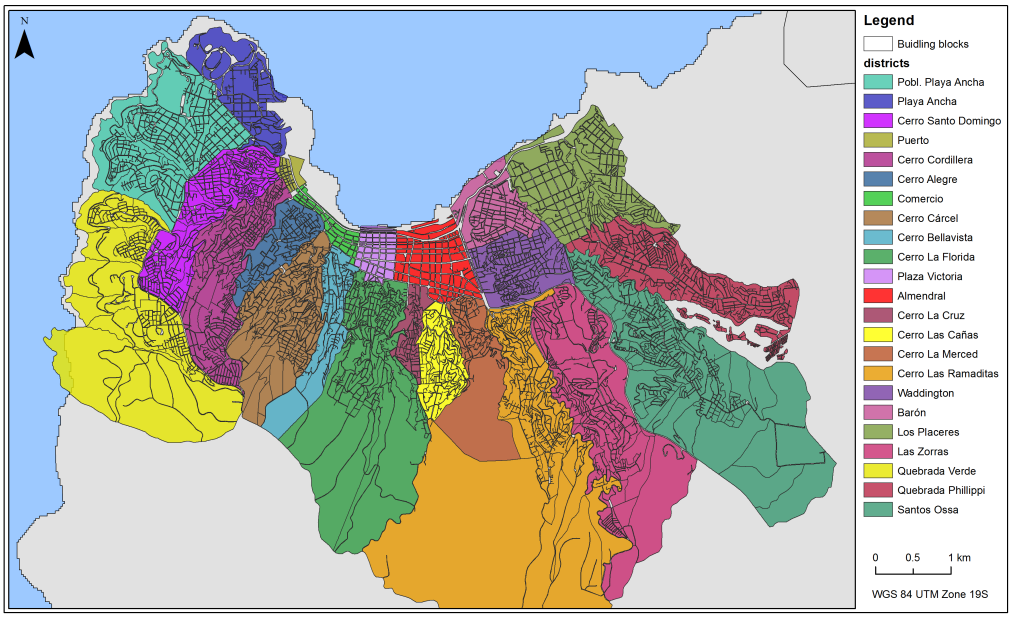


Figure 3.5.: Census data of Valparaiso 2002, 23 census districts a subdivided in 3566 building blocks

blocks). Hereby, "Comuna" refers to all inhabitants of Valparaíso, while "Distritos censales" depict subdivisions of the communal territory for the census of population and housing. They may be urban, rural or mixed. The census districts are further subdivided into census zones ("Zonas censales") which are formed by a conglomerate of building blocks ("Manzanas") to facilitate the organization, control and survey (Fig. 3.5) [Alvarado and Moya 2007].

In contrast to the census in 2012 which was a de jure census, the 2002 census was de facto. This means if the de facto residence of people is identified, this is where they happen to be on census day. The de jure residence of persons is their usual residence [Bianchini et al. 2013]. Besides the numbers of the population and households for different administrative units, the census also covers several demographic, social, cultural, educational and occupational issues.

4. Methodology & Analysis

According to the research objectives, the workflow of this study comprises four methodological steps (Fig. 4.1). Primarily, the input data requires preprocessing in order to accomplish the prerequisite to serve as base data for further analysis. Subsequently, land use and land cover (LULC) classification is performed in order to extract information from the satellite imagery. Based on the classification the number of buildings is estimated and population statistics disaggregated on pixel level in order to derive further exposure information.

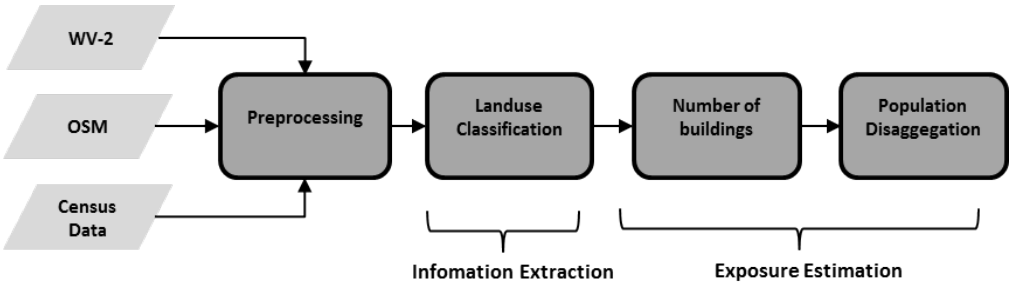


Figure 4.1.: Overwiew of methodological steps

4.1. Preprocessing

Satellite imagery, OSM-data and census data need to be preprocessed in order to meet the condition for their combination with respect to further information retrieval and analysis processes.

4.1.1. Worldview-2

The Worldview-2 scene was processed within the generic processing chain CATENA. This is an operational system for the fully automatic processing of satellite imagery developed at the German Aerospace Center [DLR 2015]. The scenes were first rectified using a physical correct orthorectification derived from the sensor model with the SRTM_CEDIT-DEM (Shuttle Radar Topography Mission), that is the SRTM C-band DEM with about 90 m resolution [Farr et al. 2007]. This leads to an absolute accurcay of 4 m. In order to

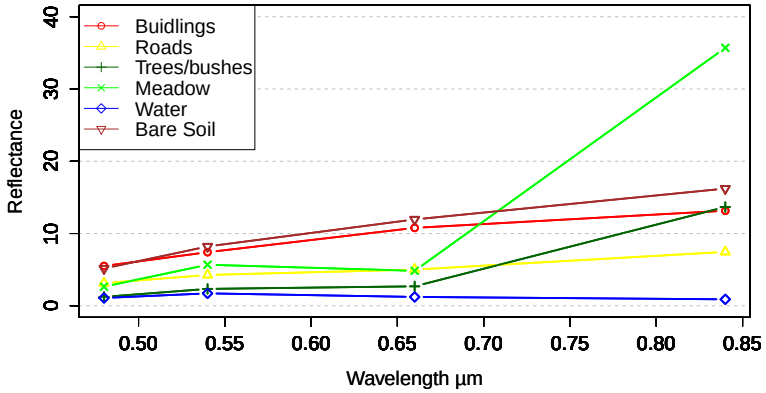


Figure 4.2.: Surface reflectance for atmospheric corrected Worldview-2 scene: Exemplary reflectance values for the landcover classes buildings, roads, trees/forest, meadow, water and bare soil.

account for radiometric distortions, atmospheric correction was conducted by means of ATCOR-2 (Atmospheric & Topographic Correction for Satellite Imagery) developed by Richter and Schläpfer [2014], which is also implemented in the processing chain. It is a physical model based on MODTRAN-4 (Moderate Resolution atmospheric transmission) for characterization of transmitting electromagnetic radiation. In other words, uncalibrated pixel values (Digital Numbers (DN)) are calibrated into physically meaningful units. Thereby, the raw image DNs are converted into Top of Atmosphere Radiance values by applying gain and offset. Afterwards, Top of Atmosphere reflectance is derived. Subsequently, surface reflectance constitutes the reflectance of the surface of the earth that is not affected by clouds and other atmospheric components [Richter and Schläpfer 2014]. The required information about sensor geometry for calibration purpose is provided in the Worldview-2 metadata. A maritime urban aerosol type was used, since the study site is located on the coast. Required atmospheric information (e.g., visibility) was estimated automatically. Further information about ATCOR is provided among others in Richter [1996]. Exemplary resulting reflectance values for six landcover classes are illustrated in figure 4.2. The provided imagery shows rather weak reflectance values, which could not be improved by atmospheric correction.

Finally, the geometric high resolution panchromatic image is combined with the lower resolved multispectral image by pansharpening in order to obtain a multispectral image of highest possible resolution. Subsequently, the resulting pansharped multispectral image features four spectral channels and a geometrical resolution of 0.5 m.

4.1.2. OpenStreetMap

For automated generation of training samples, six classes ("Buildings", "Roads", "Bare soil", "Trees/bushes", "Meadow", "Water") are distinguished and generated based on the OSM categories "buildings", "roads", "landuse" and "natural" (Tab. 4.1). Since for the study site there were no useful categories available in OSM for the class "Bare soil", several samples were digitized manually. Moreover, burnt areas during the forest fire in April 2014 were digitized and subtracted from all the training classes.

Training class	**key;value**
Buildings	buildings;yes, industrial, etc.
Roads	highway; primary, secondery, etc.
Trees/Bushes	landuse; orchard natural; forest
Meadow	landuse; meadow, grass
Water	natural; water, reservoir
Bare soil	-

Table 4.1.: Composition of OSM-based training classes: Training samples for six landuse classes are generated based on categorized OSM attributes

All OSM categories were manually checked with respect to their positional accuracy, size, shape and semantics. Low numbers of contributors and low contribution rates within the study site suggest an underdeveloped quality assurance process (Fig. 3.4). As processed objects will be deployed as training samples, they directly influence classification results and therefore outliers have to be removed from the database.
Digitized forests in OSM tend to be represented as large polygons covering big forest areas but also containing other classes like houses and streets that are located within the forest area. These overlapping classes (roads, waterways, water, bare soil, buildings, landuse) were buffered with a 2 m radius and subtracted from the forest polygons in order to obtain the actual forest class.
Concerning the samples for the class "Water", features labeled as "reservoirs" were checked whether they are covered or uncovered reservoirs in order to obtain only uncovered water bodies.
Features of the class "Meadow" were checked for trustworthiness by excluding polygons covering other additional objects besides meadow, e.g. meadows spiked with trees (e.g., cityparks).

OSM polyline features "Roads" were buffered with 1 m, resulting in a polygon feature. "steps", "pathes" and "tracks" were excluded to avoid involving materials other than concrete (e.g., soil, gravel). Especially in the hilly quarters, a lot of inaccurate streets

Figure 4.3.: Irregular offset of OSM-Buildings: Displacement (blue) has to be spatially adjusted to the satellite imagery (red).

concerning their positional accuracy and shape could be observed. Because of prevailing inaccuracy, very long and curvy streets were excluded .

Buildings also tend to have a slight nonlinear offset with respect to the Worldview-2 scene. The spatial distribution of the shifts is similar to the shifts of the roads: Buildings in the flat harbour area, in Viña del Mar and Placilla de Peñuelas are almost without any offset and size and shape are accurate. Buildings in the hilly quarters have a nonlinear offset of up to 5 m and sizes and shapes differ from related objects in the satellite scene. As stated in section 3.2.2, digitized roads and buildings in the hilly districts date back to two HOT OSM Tasks in April 2014. As base imagery for the mapping Bing imagery was used here [OSM Task Manager 2014b;a, Global Resources News 2014]. Bing imagery is projected in the rather inaccurate Web Mercator Projection which leads particularly in the study area to skewed camera angle. Concerning the off-nadir acquisition, it depends on the contributor whether he/she maps on basis of rooftops or ground plan of buildings, respectively. To minimize possible error sources, high buildings like skyscrapers were therefore excluded from the samples.

Additionally, the alignment of Bing imagery is inconsistent across zoom levels [OSM–Wiki 2015]. Areas with height differences (e.g., containing mountains, hills, bridges) seem more prone to this effect, which suggests that the base imagery was not always taken at a precise 90 degree angle. As a result, the geometric accuracy of objects at the same height (ground level in a flat area) may be quite precise while objects which are higher (embankments, bridges, buildings) or lower (cuttings) may be slightly off. This could explain the fact that geometric inconsistencies between Bing and Worldview-2 appear especially in

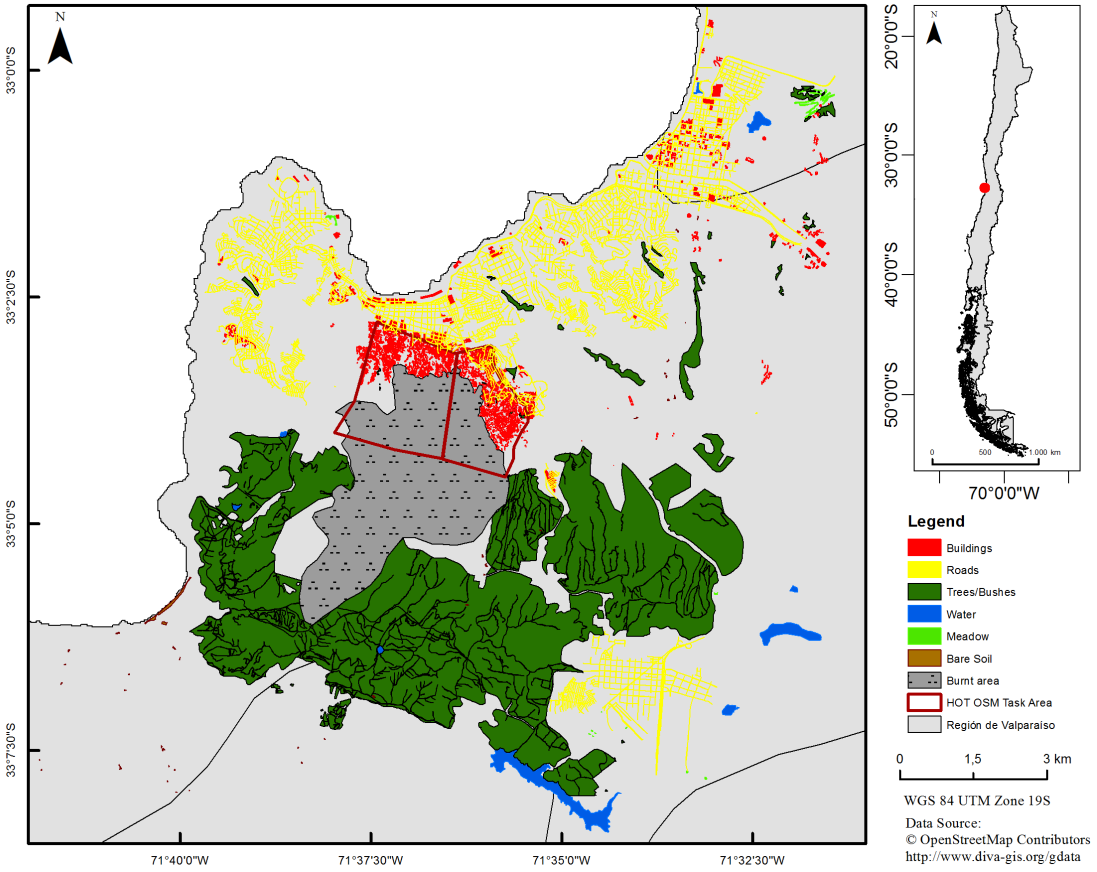


Figure 4.4.: Overview of underlying OSM data: Left: Extent of OSM input dataset classified in six landuse classes. HOT-OSM area is marked in red. Right: Localisation of study site in Chile.

hilly regions.

Moreover, the time of acquisition of the available Bing imageries varies between the zoom levels, whereby imagery at the lowest zoom levels date back to November 2013 [OSM 2015b;a]. However, offsets appear also between Bing imagery and OSM, which suggests that the Bing imagery available for mapping in April 2014 differs from the current available ones. The OpenStreetMap community is aware of alignment difficulties and therefore suggests to utilize the "Imagery Offset database Plugin" for the use within the java-based extensible editor for OSM (JOSM) in order to correct emerging offsets with the help of ancillary data like GPS tracks [OSM-Wiki 2015]. However, for the study site, no offsets neither useful GPS tracks were registered in the database. Consequently, spatial adjustment for buildings had to be conducted manually by georeferencing the polygons to the satellite image. Yet, it has to be considered that reasons for possible offsets can also arise from the Worldview-2 scenes, since the orthorectification has an absolute accuracy of 4 m and therefore derivations are possible (Sec. 4.1.1). Figure 4.4 illustrates the preprocessed

OSM input dataset.

4.1.3. Census Data

Census data of Valparaíso have to be extracted from the census database provided by the National Statistics Institute of Chile (INE) utilizing REDATAM (Retrieval of Data for small Areas by Microcomputer). Subsequently, by means of data processing the census data is available in tabular form. Associated spatial units to locate census data geographically are available as shapefiles. The tables have to be arranged due to their structure in order to be able to join them with the shapefiles in a second step. This is achieved automatically with the help of a python script. The spatial dataset reveals irregular offsets and has to be spatially adjusted. Therefore it is georeferenced to the satellite image. Finally, census information is joined with the associated spatial units.

4.2. Information Extraction

A first crucial step in estimating exposure information from very high-resolution satellite imagery (VHRS) is the extraction of a detailed built-up mask that provides information about the location of buildings and the area occupied by buildings [Wieland et al. 2012b]. Therefore, a detailed LULC map is derived. For this purpose a hierarchical supervised classification approach is designed. In the scope of object-based image analysis, the

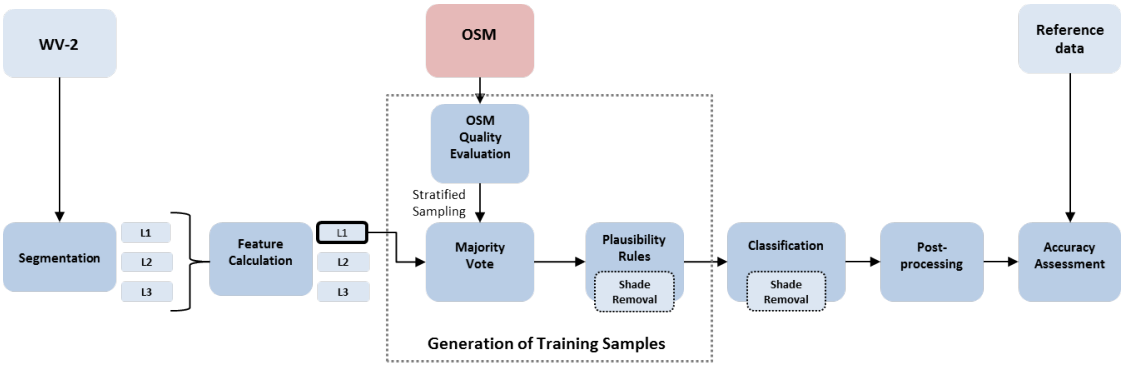


Figure 4.5.: Scheme of the proposed Information Extraction Workflow: Satellite Imagery (WV-2) is hierarchical segmented and descripitve features are calculated for three resulting segmentation levels (L1, L2, L3). In combination with the OSM dataset, training segments are generated based on L1 in order to train the classifier. Preprocessing is applied on the classification result and evaluated in terms of accuracy assessment.

satellite image is subject to a multiscale segmentation in a first step and features are computed for the resulting image objects in a second step. Then, the first segmentation level is fused with the OSM data in order to create training samples. By considering OSM data quality and plausibility rules, best possible reliability of samples is ensured. Shade

is excluded from the samples before the training and after the classification phase, what renders the approach hierarchical. Finally, classification postprocessing and accuracy assessment are conducted (Fig. 4.5). The single work steps will be presented in the following section.
Due to the size of the scene which would induce processing times of several days, the scene is divided in six subsets. The borders of the subsets were generated in accordance to natural transitions of landcover (i.e., forested areas, rivers, etc, see Appendix A.1). Each part is processed separately to optimize processing times.

4.2.1. Segmentation

The first step of an object-based classification approach is generating meaningful image objects (IO) by image segmentation [Baatz and Schäpe 2000]. The multiresolution segmentation approach which is implemented in Definiens eCognition Developer by Trimble GmbH [2014] is applied. It is a bottom-up region-merging technique, hence a region-based algorithm, that locally minimizes the average heterogeneity of image objects for a given resolution [Baatz and Schäpe 2000]. In this context, the algorithm starts considering each pixel as a seperate IO. Then, they are merged to bigger objects based on local homogeneity criterion, describing the similarity between adjacent IOs [Darwish et al. 2003]. This process is controlled by a user-defined scale, as well as color/shape and smoothness/compactness weights [Johnson and Xie 2013]. The scale parameter is a function of geometric resolution of the image data and defines maximum allowed heterogeneity for the resulting segments. The homogeneity criterion is composed of color and shape (smoothness and compactness) [Trimble GmbH 2014].
Urban structures and objects feature several magnitudes of spatial extent. This limits the use of a single segmentation scale and gives rise to a multilevel, hierarchical representation of the objects in the image domain [Baatz and Schäpe 2000, Li et al. 2011]. Multiscale segmentation procedures have been proposed in literature, especially for complex VHR imagery, where thin, enveloped and nested regions must be retained [Bruzzone and Carlin 2006, Esch et al. 2008, Pesaresi and Benediktsson 2001]. Here, the use of a single scale factor for a whole scene would lead to over-segmentation (one real world object contains several IOs) or under-segmentation (one IO encloses several real-world objects), whereas the latter one is least desirable and leads to errors [De Roeck et al. 2009].
The multiscale approach establishes precise hierarchical relationships between each pixel in the image and the region that adaptively defines its context at different levels [Bruzzone and Carlin 2006]. Consequently, the aim is not to recognize the best level of representation of each object like in traditional segmentation approaches, but rather to model the multilevel spatial context of each pixel. Thus, IOs generated at fine scales are nested

inside of IOs generated at coarser scales. Larger objects are referred to as super-objects of the smaller segments, and the smaller segments are referred to as sub-objects of the larger objects [Johnson and Xie 2013].
In this study, segmentation is carried out on three levels (L1, L2, L3) with scale factors of 50 (L1), 150 (L2) and 500 (L3). From a conceptual point of view, the smallest level (L1) comprises objects smaller than real-world objects (oversegmentation), on L2 generated objects correspond most real-world objects of interest (buildings) and L3 comprises objects bigger than most real-world objects in the image (undersegmentation).
During the process, constant weights for color and shape criteria are maintained throughout the segmentation levels. Shape is determined with a factor of 0.7, which emphasizes shape heterogeneity rather than the influence of color throughout the segmentation process. This is due to the fact that man-made features have characteristic shape and size properties and very heterogeneous spectral characteristics [Dahiya and Garg 2013]. The output of the approach is a hierarchy of regions in contrast to a single partition [Li et al. 2011].

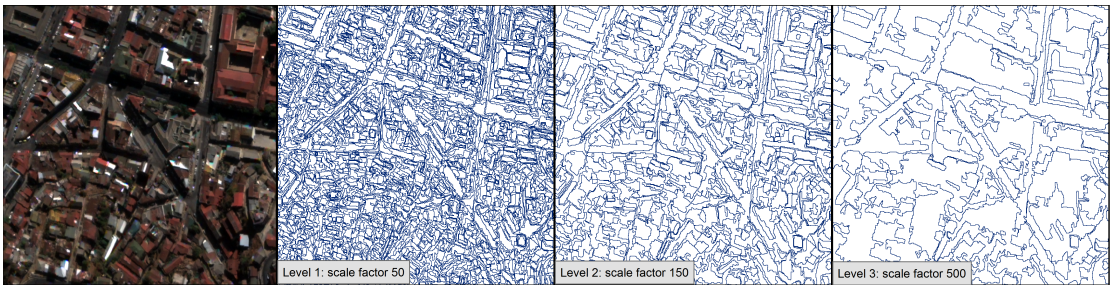


Figure 4.6.: Multiscale Segmentation: Hierarchical segmentation on three levels. Level 1 Objects: Oversegmentation with scale factor 50, Level 2 Super-Objects: Representation of real-world objects (in particular buildings) with scale factor 150, Level 3 Super-Objects II: Undersegmentation with scale factor 500.

4.2.2. Feature Calculation

The exclusive use of spectral information is often not sufficient for VHR urban image classification. Therefore, spectral-spatial approaches have been designed. Features are used to describe characteristics of IOs in order to differentiate them as good as possible against each other in the classification phase. For this purpose, a wide range of features is derived in order to represent heterogeneities and homogeneities between IOs in a wide range. The feature calculation is carried out consistently for three segmentation levels (see Sec. 4.2.1, Fig. 4.6). The inclusion of information from coarser segmentation scales allows to model spatial context information and relations in feature space adequately [Johnson and Xie 2013, Bruzzone and Carlin 2006]. Thereby, the IOs of the smallest

level (L1) will not only hold information about their own characteristics, but also the information of their two super-objects, respectively. The level 1 dataset, comprising all available information, will be the base for further analysis. Computed features are listed in table 4.2 and will be outlined in the following.

In the context of image analysis techniques, mathematical morphology has been proven useful for the classification of land use in urban scenes [Acar and Bayram 2009, Singh et al. 2012, Tuia et al. 2009, Pierre Soille 2003]. It is especially successful if image objects are characterized by clear shape, size, and contrast properties [Giada et al. 2003, Pierre Soille 2003]. Particularly urban areas are prone to noise and illumination problems due to heterogeneous structures and different heights, and can therefore profit from morphological filtering.

The operators Erosion, Dilation, Opening, Closing, Opening by Top Hat and Closing by top hat [Pierre Soille 2003] were implemented exclusively for this study in the eCognition software. Thereby, the choice of the size of the structuring element (SE) is crucial [Pesaresi and Benediktsson 2001, Pierre Soille 2003]. Shape and size of the SE should be chosen according to the image resolution, the size of the image structures to be extracted and their main direction in the image [Tuia et al. 2009, Giada et al. 2003, Pierre Soille 2003]. It is preferable to use a range of SE sizes, as some structures may have a high response to a given SE size and a lower response for other SE sizes [Pesaresi and Benediktsson 2001]. To ensure liable performance of the morphological operators, a square-shaped SE with ascending sizes in the range SE $\in [5, 10, 15, 20, 25, 30]$ was deployed. Herewith it is attempted to cover several sizes of different buildings.

Erosion and Dilation are the elementary operators in mathematical morphology on which all other transformations are based. Erosion assigns each pixel the minimum value of the pixels inside the SE respectively, and thus removes not only all components that cannot contain the SE, but shrinks all other components [Giada et al. 2003]. On the contrary, dilation calculates the maximum image value for each pixel within the neighborhood defined by the SE [Giada et al. 2003]. Thus, erosion shrinks an image according to the shape of the SE, and dilation grows the objects by the same shape.

Information from the original image that is lost during erosion, is tried to be recovered by means of the operator morphological opening by dilating the eroded image [Giada et al. 2003, Pierre Soille 2003]. In terms of an urban environment, opening can for example achieve to remove trees attached to buildings [Awrangjeb et al. 2011].

In contrast, morphological closing is achieved by eroding the dilated image, recovering the initial shape of the image structures that have been dilated [Pierre Soille 2003]. Opening and Closing operators erase structures that are smaller than the SE. If a grayscale image is interpreted as a topographical relief, then the opening cuts peaks and closing fills valleys that are smaller than the SE [Pesaresi and Benediktsson 2001]. This effect can be

observed by computing the residual between the filtered image and the original image in the top hat transform. The operators Opening by top hat and Closing by top hat take the residuals of the Opening/Closing and original images to isolate structures that are smaller than the given SE.

Furthermore, the mean and the median of the spectral channels, the brightness and the maximum difference (Max.Diff) are derived as measures of central tendency. The brightness depicts a value over all image layers taking into account their darkest and brightest possible intensity. In addition, Max.Diff represents the difference of the mean intensities relative to the brightness [Stumpf and Kerle 2011].

Measures of spread describe the similarity or variability of values per segment and are measures of statistical dispersion. Thereby, standard deviation, variance, minimum, maximum, range and interquartile range (IQR) were derived for the spectral bands Blue, Green, Red and NIR (Table 4.2). Variance is a measure of the degree of spread around the mean value. The standard deviation is determined as the square root of the variance. Both measures reveal the degree of noise within an object. While the range is the difference between the lowest and highest values (maximum and minimum), the IQR is the difference of the upper and lower quartiles (Q25,Q75). Since the range is solely based on the most extreme values within an object (often outliers) it can result in a range that is not typical of the variability within an IO. In contrast, the IQR indicates the extent to which the central 50% of values within the object are dispersed and therefore can be regarded as less prone to outliers.

Geometrical features describe the extent and shape of the IOs. They are particularly important as super-object information (e.g., long streets, building blocks). Segments of the smallest scale do not show specific geometric patterns throughout different classes since real-world objects comprise several segments. This is due to great structural differences within one real-world object because of the very high resolution satellite image. The area of an IO is described by the number of pixels forming it. The length of an IO describes the length of the polygon's main line while the width describes the width of the polygon's main line. In addition, the border length is defined as the sum of the edges shared with other IOs [Martha et al. 2010].

Moreover, the shape of the object can be described by several features. The asymmetry is described by the relative length of an IO compared to a regular polygon, which is delineated by approximating an ellipse around the respective IO. The more longish an image object is, the more asymmetric is its shape. Furthermore, the compactness of an IO describes how compact it is based on its area. In contrast, the density feature describes the distribution of pixels of an IO in space. The more its shape resembles a filament rather than a square, the lower its density [Moine et al. 2009]. Similarly, the epileptic fit, as well as the roundness feature, characterize the fit of an object into an ellipsoid, but

Morphological Operators	Erosion (Pan, SE=5,10,15,20,25,30) Dilation (Pan,SE=5,10,15,20,25,30) Opening (Pan, SE=5,10,15,20,25,30) Closing (Pan, SE=5,10,15,20,25,30) Opening by Top Hat (Pan, SE=5,10,15,20,25,30) Closing by Top Hat (Pan, SE=5,10,15,20,25,30)
Measures of central tendency	Mean (ρB, ρG, ρR, ρNIR, ρBrightness, ρMax.Diff) Median (ρB, ρG, ρR, ρNIR, ρBrightness, ρMax.Diff)
Measures of spread	Standard Deviation (ρB, ρG, ρR, ρNIR) Variance (ρB, ρG, ρR, ρNIR) Minimum (ρB, ρG, ρR, ρNIR) Maximum (ρB, ρG, ρR, ρNIR) Range (ρB, ρG, ρR, ρNIR) Interquartile Range (ρB, ρG, ρR, ρNIR)
Geometry	**Extent**
	Area Border Length Length Length/Width
	Shape
	Asymmetry Compactness Density Rectangular Fit Elliptic Fit Roundness Shape Index Perimeter
Contextual measures	Relative border to shade Number of sub-objects Deviation sub to super-object (ρBrightness)
Ratios	Normalized Differenced Vegetation Index Ratio Red/NIR Ratio Area/Perimeter

Table 4.2.: Feature computation: Each IO is represented by a 270-dimensional feature vector, containing 90 different features computed for three segmentation levels. Note that morphological operators are computed for the range of SE sizes while measures of central tendency and spread are computed for each listed channel. All computations are conducted in eCognition Developer. ρ = reflectance

calculated in different ways [Moine et al. 2009]. In contrast, rectangular fit approximates the IO by a rectangle of similar size and proportions. Finally, the perimeter is the sum of all edges of an IO, while the smoothness of an IO border is characterized by the shape index. The smoother the border the lower the index [Martha et al. 2010].
Contextual features make use of class relations, neighborhood or sub- and super-object information. The relative border to a shade segment is computed for all three levels. Buildings often cast shadows on streets or other buildings due to relatively high altitudes. The introduction of those features is motivated by the fact that buildings most likely appear in groups separated by streets. Thereby, the relative border to shade could be useful to identify a built-up environment.
Concerning sub- and super-object relations, it is suggested that super-objects containing high numbers of sub-objects represent complex structures, which are most likely urban structures and primary building blocks of houses. Furthermore, high difference rates of brightness values between sub- and super-objects indicate changes of structures within a super-object. Changes possibly represent differing objects within an super-object, which could be cars on roads, a building in an vegetated area etc. (see Table 4.2). Those noisy areas emphasize complex structures which are often rather man-made than natural phenomenons.
Last but not least, the Normalized Differenced Vegetation Index (NDVI), Red/NIR ratio and the Area/Perimeter ratio are calulated. In particular the NDVI is useful in order to isolate natural regions from manmade regions [Albertz 2013, Singh et al. 2012]. Moreover, in terms of the Red/NIR ratio specific building types can be highlighted. The Area/perimeter ratio is another measure of IO complexity.
Finally, the Level 1 segmentation comprising all calculated features based on all segmentation levels is exported as shapefile. Overall, each IO is represented by a 270-dimensional feature vector, containing 90 different features computed for three segmentation levels.

4.2.3. Generating Training Samples from OSM Data

Until now, segmentation level 1 comprising the IOs described by 270 features and the OSM data are available separately (see Fig. 4.7, a and b). In a next step the two datasets have to be combined in order to obtain training samples based on the segmentation. Therefore, a majority vote is introduced for the IOs of the first segmentation level: An IO is classified as a training sample with respect to a certain thematic class if more than 50% of the area overlaps with the OSM-data of the same class. Figure 4.7 illustrates the fusion of the two datasets. After their combination by majority vote (Fig. 4.7, c), the credibility of the dataset will be ensured by means of the introduction of plausibility rules (Fig. 4.7, d).

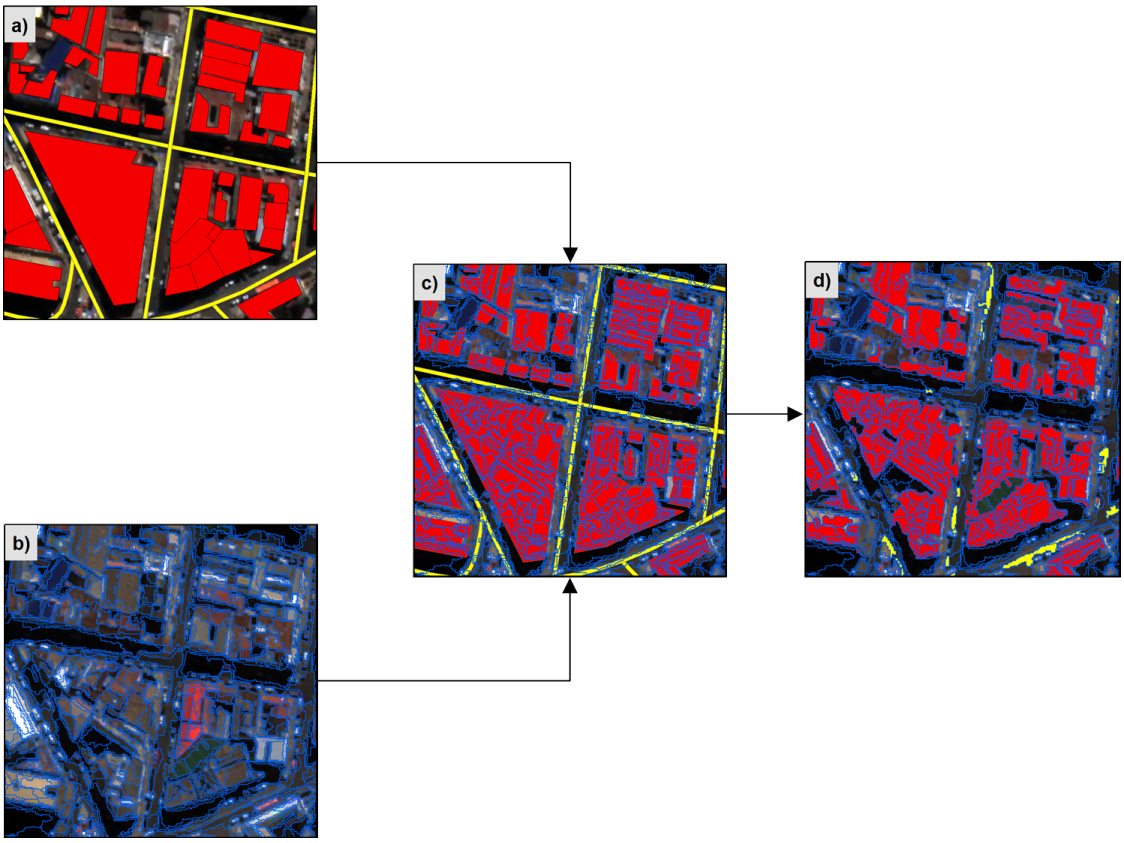


Figure 4.7.: Generating training segments from OSM data. a) OSM data, b) Image segmentation, c) Fusion of OSM and IOs by majority vote, d) Generated training segments based on Plausibility rules (Fig. 4.8)

However, as a first step of the hierarchical classification approach, shade needs to be excluded from the generated training and validation datasets by the brightness value of the IOs (Fig. 4.8). Especially in urban areas, shadows cause major difficulties to obtain useful landuse maps, because information in shady areas is distorted or lost [Van de Voorde et al. 2007]. Shade is not included in the training classes, because it is not an available attribute in OSM-data. Due to differing data acquisition with varying sun azimuth angles the introduction of the class shade into OSM is not suggestive. Since specifically dark objects refer to shade or water, water samples are precluded from this rule. Later on, the same rule will be applied on the classification results in order to label shade.

As shown in figure 4.8, several plausibility rules are introduced to increase credibility of the OSM-based training samples and to prevent corrupt training input. The aim is to obtain training samples as accurate as possible. This is highly relevant due to the fact that the classifier will be learned based on the given training set. Plausibility checks are

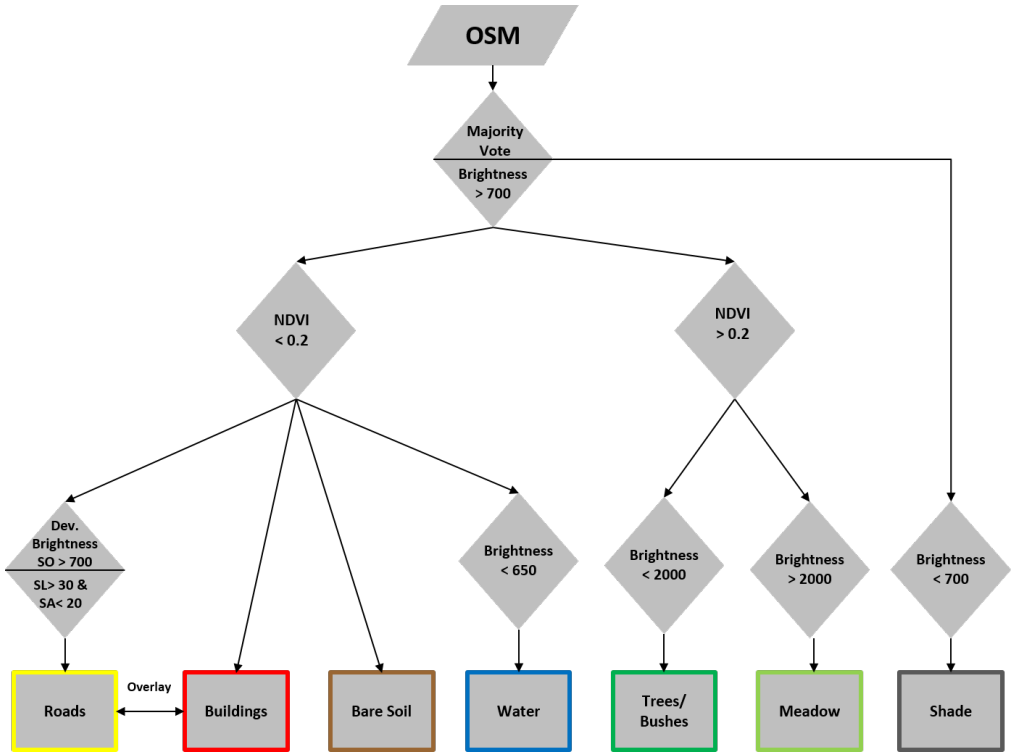


Figure 4.8.: Plausibility rules for training samples. All classes: all segments have to have an overlap of at least 50% with one OSM class (majority vote) , exclusion of shade (dark segments (Brightness<700)) that are not water), **"Buildings"**: exclusion of adjacent vegetation, overlaid by roads; **"Roads"**: exclusion of roads covered by trees (alleys), overlaid by buildings and noise (mainly cars); **"Trees/bushes"**: appropriate high NDVI; **"Meadow"**: appropriate high NDVI and dissociation of darker forests; **"Water"**: exclusion from nearby vegetation by appropriate lower NDVI, low brightness for exclusion of adjacent beaches, islands, ships; **"Bare soil"**: dissociation of vegetation. SO = Super-Object, SL = Shape Length, SA = Shape Area, NDVI = Normalized Differenced Vegetation Index.

mainly based on NDVI and brightness values since these features are easy to interpret and are capable of distinguishing roughly between certain classes. The rules are carefully chosen and systematically tested by analyzing several areas of the study site.
In this context, figure 4.8 illustrates the hierarchical model consisting of several binary decisions with formulated conditions the data has to meet in order to be assigned to its target class. Thereby, the class "Buildings" requires a NDVI value smaller than 0.2. In this way, adjacent vegetation can be excluded from the dataset. The same rule applies to the class "Roads", preventing the inclusion of road objects that are actually covered by trees. Roads are also often subject to noise (e.g., cars) and training samples comprising noisy elements could distort the class representation. Therefore, a rule is introduced in order to identify small noisy objects based on their brightness deviation to the respective super-object as well as their size (Fig. 4.8). Furthermore, both classes are subtracted from each other in order to prevent a mutually overlay that could possibly arise from the

fusion of the dataset by majority vote. The classes "Trees/Bushes" as well as "Meadow" should exceed a certain NDVI value (Fig. 4.8), whereas both classes can be distinguished by their brightness (the class "Meadow" reveals higher brightness values than the class "Trees/bushes"). The class "Water" can be distinguished from nearby vegetation by a lower NDVI value. Moreover, adjacent (bright) beaches as well as islands and ships can be excluded from the water samples by a low brightness value. Finally, the class "Bare soil" is restricted by a small NDVI value.

Assessing OSM Quality for Stratified Selection of Training Samples

In order to ensure best possible quality training input, the OSM quality is further assessed and samples subsequently selected with regard to their reliability. Since there is no reference data to assess the quality of the input OSM data, the behavior of the community can serve as a quality indicator in this study. Arsanjani et al. [2015] analyzed spatio-temporal patterns of contributions to OSM by proposing a Contribution Index (CI) based on the OSM history, which can serve also in this study as an intrinsic quality measure. The CI is a function of four variables, namely "Quantity", "Interactivity", "Semantics" and "Attractivity" that are derived based on a grid.
"Quantity" counts the number of nodes given in each grid cell, whereas "Interactivity" determines how many times the nodes within each cell have been edited by averaging the number of versions ("osmversion"). "Attractivity" quantifies the number of users that have edited nodes within each cell, representing the ability of the cells to attract contributors. The number of nodes in each cell that are given attributes are indicated by the variable "Semantics", showing how well the nodes within a grid cell are attributed [Arsanjani et al. 2015].
OSM history data is extracted from the planet file and analyzed based on 1 km^2 grids in terms of the 4 variables as suggested by Arsanjani et al. [2015] (Fig. 4.9). Unlike proposed by Arsanjani et al. [2015], the overall index is calculated by introducing a simple school grading system, where a cell gains five points for the best possible class and one point for the lowest possible class, respectively. Then, the average grade per cell is calculated, resulting in a value range where one would signify that the worst class was gained in all four variables and 5 would signify the gain of the highest class in all four variables.
The CI result represents what is already suggested by the result of the single variables, respectively. The urban area performs better than the rest of the study site in terms of the contribution index, which complements the findings of Arsanjani et al. [2015] and Zielstra and Zipf [2010]. The resulting CI matrix gives an idea of the spatial distribution of quality within the analyzed area and can be considered for prioritized selection of training samples.

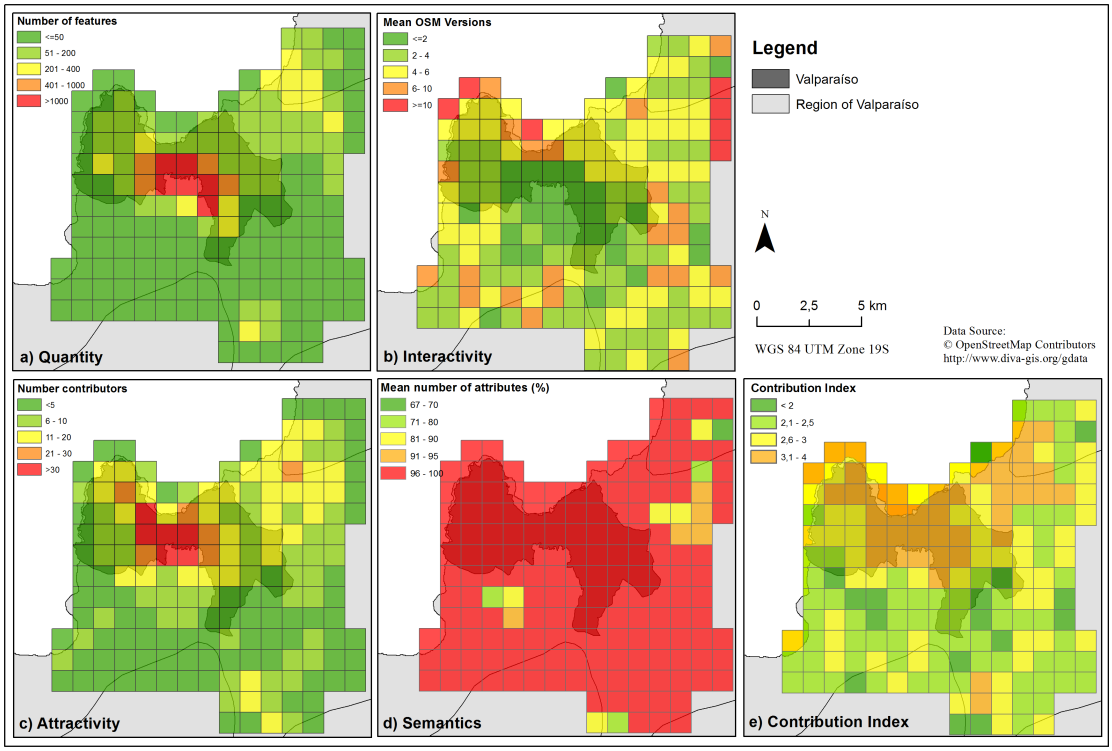


Figure 4.9.: Contribution Index (CI) proposed by Arsanjani et al. [2015] in order to assess OSM quality: The CI consists of **Quantity (a)**: Number of nodes per cell, **Interactivity (b)**: Average number of edits made per cell , **Attractivity (c)**: Number of users that edited nodes within a cell, and **Semantics (d)**: Mean number of nodes that posess attributes per cell.

Available OSM samples per class do not implicitly correlate to the occurrence of the respective class in reality. Thus, in order to avoid model bias by imbalanced data, the same number of samples should be selected randomly for every class from the whole training population throughout the scene. Since the samples are selected from a big population they should depict a well representation of the base training data. The number of samples per class is determined by available samples for the smallest class which results in 600 samples per class. Since the CI highlights areas of better quality of OSM data throughout the study site, we design that more training samples should originate from this area. Yet, not all samples should originate from the highly spatially aggregated CI quality area in order to maintain the distribution throughout the whole scene and avoid possible problems associated to covariance shift. 2/3 of the samples per class are therefore selected randomly from CI quality areas whereas 1/3 is taken randomly from the residual areas. For smaller classes (e.g, water) existing samples in quality areas are captured, if their amount is not sufficient to reach 2/3, lacking samples are simply randomly taken from residual areas.

4.2.4. Classification Model Training

The second working step of OBIA comprises the classification of the image objects derived in the segmentation process. In this study, a relatively new classifier ensemble method called Rotation Forest (RTF) proposed by Rodriguez et al. [2006] is applied. Classifier ensembles have been successfully used in recent years and are especially suggested for high-dimensional data [Waske et al. 2009]. Especially the Rotation Forest algorithm has shown promising results and has outperformed most widely used ensemble methods, such as Random Forest, Classification and Regression Trees, Bagging, etc. [Kavzoglu and Colkesen 2013, Kuncheva and Rodriguez 2007, Rodriguez et al. 2006, Stiglic et al. 2011, Xia et al. 2014].
RTF is a non-parametric, supervised classifier ensemble that uses decision trees as base classifiers and conducts feature extraction (i.e., feature axis rotation) based on principal component analysis (PCA) [Rodriguez et al. 2006]. Thereby, RTF is based on the sensibility of decision trees to axis rotations. Classifiers obtained with different rotations of a dataset can be very diverse [Stiglic et al. 2011].

It is assumed that there is a dataset containing the training objects in the form of an $N * n$ matrix, with N data points $\chi = [x_1, ..., x_n]^T$ described by n features (in this study $n = 270$). $Y = [y_1, ..., y_n]^T$ would be a set of class labels for the data, where y_i takes a value from the set of class values $\omega_1, ..., \omega_c$. The several classifiers in the ensemble are referred to as $D_1, ..., D_L$ and the feature set as F [Rodriguez et al. 2006]. Thereby, the ensemble size L has to be chosen in advance [Kavzoglu and Colkesen 2013]. To construct a training set for a classifier D_i several steps are carried out (see Fig. 4.10):
Initially, RTF randomly splits the feature set F into K feature subsets ($F_{i,j}$, $j = 1...k$) [Rodriguez et al. 2006]. Subsequently, each split will be formed by only one attribute of the dataset and each attribute of a new dataset corresponds to only one group (group = number of features in each subset). Each subset ($F_{i,j}$) contains $M = n/K$ features ($X_{i,j}$). From $X_{i,j}$ a new training set $X'_{i,j}$ is selected randomly with 75% size using a bootstrap algorithm.
PCA is then run separately on each subset $X'_{i,j}$ to obtain the coefficients $a_{i,2}, ..., a_{i,2}$ in a matrix $C_{i,j}$. However, PCA is not used to reduce the dimensionality, but only to rotate the original feature axes in order to obtain different training sets for learning base classifiers [Stiglic et al. 2011, Lasota et al. 2012]. On base of the obtained coefficients a rotation matrix R_i is organized whose columns are then rearranged to R_i^a with respect to the original feature set (Fig. 4.10). In other words, the data is transformed linearly into the new feature space [Gaikwad and Pise 2014]. Thus, a new extracted training set XR_i^a, is reassembled while preserving all components. It is used for training each decision tree classifier respectively, whereby all classifiers are trained in a parallel manner [Xia et al.

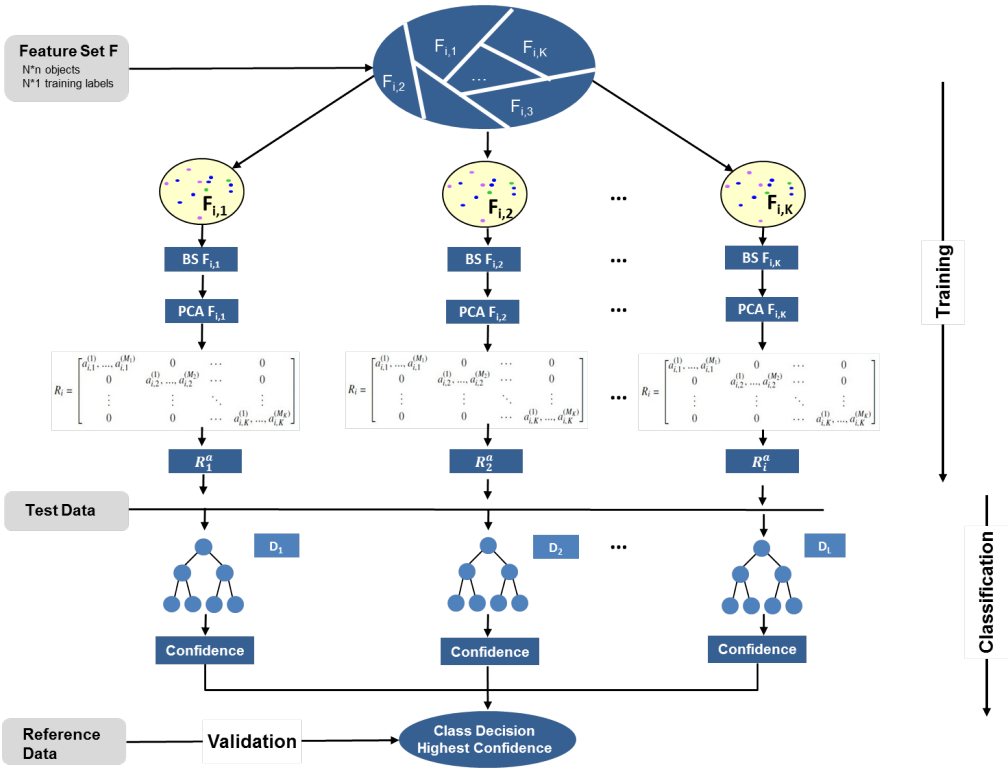


Figure 4.10.: Rotation Forest ensemble method proposed by Rodriguez et al. [2006]: The training feature set F is split into K subsets and transformation algorithms (BS = Bootstrap, PCA = Principal Component Analysis) are run separately on each subset, reassembling a new extracted feature set while keeping all the components. The approach promotes the generation of accurate and diverse classifiers. Classification was carried out within the Waikoto environment for knowledge analysis software (WEKA), which is an open source java-based software developed for machine learning techniques and data mining processes [Machine Learning Group at the University of Waikato 2015]

2014]. Hereby, the selection of a sample size $X'_{i,j}$ smaller than $X_{i,j}$ by bootstrap sampling, and thus running PCA on a subset of classes rather than the whole dataset, has two reasons: to avoid identical coefficients if the same feature subset is chosen for different classifiers and to enhance diversity among the generated ensemble member classifiers [Rodriguez et al. 2006].

Finally, for a given test sample x, the confidence is calculated for each class by the average combination method as shown in equation (4.1), where L is the number of classifiers in the ensemble and $d_{i,j}(xR_i^a)$ is the probability assigned by the classifier D_i to the hypothesis that x comes from the class ω_y [Rodriguez et al. 2006].

$$\mu_j(x) = \frac{1}{L}\sum_{i=1}^{L} d_{i,j}(xR_i^a), j = 1, ..., c. \tag{4.1}$$

x will be assigned to the class with the highest confidence.

The success of the rotation forest method lies in the application of the rotation matrix constructed by linearly transformed subsets. Thus, an effective transformation method has significant impact on the classification accuracy [Kavzoglu and Colkesen 2013]. However, the practical application of the RTF requires some parameters to be fixed in the model construction stage. Here, they were chosen carefully based on literature review and systematic experiments. According to Xia et al. [2014] the sensitive parameters are the ensemble size L and the number of features in the subset, the latter one is also referred to as indicator of the operating complexity [Rodriguez et al. 2006]. Therefore, these two parameters need to be fixed in order to optimize the classification accuracy.
The J48 classifier from the WEKA library is selected as base classifier to construct the trees. It is a reimplementation of C4.5 decision tree and uses gain ratio as splitting criterion. The J48 classifier uses an error-based pruning algorithm. The size of decision trees can be reduced by pruning, removing sections of the tree that provide little power to classify instances. Thereby too complex trees as well as the effect of overfitting are avoided [Duda et al. 2001]. The parameter "Instances to be removed" is set to 22 due to highest result accuracy. Thus, the confidence value used when pruning the tree is set to 25% as proposed by Rodriguez et al. [2006].
The ensemble size L can be regarded as a hyper parameter of the ensemble and can be tuned by the seperate validation set. The classifier model was tested for a varying number of iterations between 1 and 100, with 10 iterations showing best results.
PCA is used as feature transformation algorithm as suggested by Rodriguez et al. [2006], Kuncheva and Rodriguez [2007] and Xia et al. [2014]. The number of splits (size of groups) that control the sizes of the feature subsets was set to three features per subset as suggested by Rodriguez et al. [2006] and Kuncheva and Rodriguez [2007].

Due to data size the scene is divided into 120 subsets and the classifier model is applied on each. Due to the hierarchical classification approach (see Sec. 4.2, Fig. 4.5) shade classification is accomplished in an independent step. As stated in section 4.2.3, shade is already excluded from the training data, subsequently it has to be reclassified in a post-processing step to avoid the confusion with other classes, especially water [Van de Voorde et al. 2007]. It is classified by the same brightness threshold as in section 4.2.3 (also see Fig. 4.8). Once again, the exclusion of valid water classification is prevented.

Selection of Reference-Instances for Accuracy Assessment

In order to evaluate the performance of the classification model, a reference dataset is necessary. The OSM dataset is not proven to be correct, and therefore no cross validation should be performed based on it. The reference dataset is generated manually: The refer-

ence instances used to create the confusion matrix are selected following a simple random sampling. Simple random sampling implies that each sample unit (segments/image objects) in the study area has an equal chance of being selected [Congalton and Green 1999]. Randomness ensures that all portions of a region are equally likely to be selected from the sample, thereby yielding data that accurately represent the area examined and satisfying one of the fundamental requirements of inferential statistics [Campbell and Wynne 2011]. The number of observations determines the confidence interval of an estimate of the accuracy of a classification. A large sample size decreases the width of the confidence interval. Not only observations for the entire image have to be considered, but also those that are allocated to each class. It is necessary to have some minimum number of observations assigned to each class [Campbell and Wynne 2011]. Thus 0.004 % of the segments per

	Tile 1	**Tile 2**	**Tile 3**	**Tile 4**	**Tile 5**	**Tile 6**	**Sum**
Buildings	206	55	5	68	-	-	334
Roads	44	22	2	23	2	1	94
Trees/Bushes	58	46	75	124	47	49	399
Meadow	4	42	19	-	-	4	69
Water	-	13	18	-	-	4	35
Bare soil	17	27	19	45	-	17	125
Sum	329	205	138	260	49	75	**1056**

Table 4.3.: Reference instances: Random selection of 0.004 % of segments throughout six subsets of the scene (see chapter A.1), resulting in 1056 reference instances. Reference instances are manually labeled based on visual interpretation.

subset were randomly selected using ArcGIS software. Related to the number of segments per subset, this amount results in about 50-330 segments for each of the six subsets. The selected reference-instances are then manually labeled based on the visual interpretation of the input image. Table 4.3 shows the distribution of the 1056 test-instances. Instances of class water were rare because almost all available water samples were already labeled as training samples. However, a simple random sample of sufficient size clearly affords the most robust inference of map accuracy [Campbell and Wynne 2011].

4.2.5. Classification Postprocessing

To further enhance the classification accuracy, Classification Postprocessing (CPP) was carried out. Confusions in the classification results often arise from small regions with arbitrary shapes that are produced by image segmentation and are then misclassified. This results in a noisy classification map [Zheng et al. 2014].
According to Huang et al. [2014], CPP is "a refinement of the labeling in a classified image in order to enhance its original classification accuracy". This is usually achieved by the application of context-based rules or filters to the classification result [Van de

Voorde et al. 2007]. Candidates for post-classification are especially confused classes and classes that are of interest for further investigations. It is intended to improve the discrimination between pairs of classes and thereby enhance the overall accuracy by focusing on the weakest aspects of the classification [Campbell and Wynne 2011]. The class "Buildings" is not only of importance for the subsequent exposure estimation analysis but also problematical to derive, especially in a complex urban structure of VHR imagery like the study site (see Sec. 2.2). Therefore, it represents the primary object of interest for post-classification.

In Trimble's eCognition software, several context-based rules are implemented in order to correctly classify incorrect outliers and thus "cleaning up" the classification result. The use of post-classification techniques requires caution since parts of the study area might improve while others diminish by the introduction of a rule. While defining rules, care was taken to ensure that the rule set could be successfully applied under various morphological conditions, since the classification result should not be sophisticated. For this reason only simple and unambiguous rules are introduced.

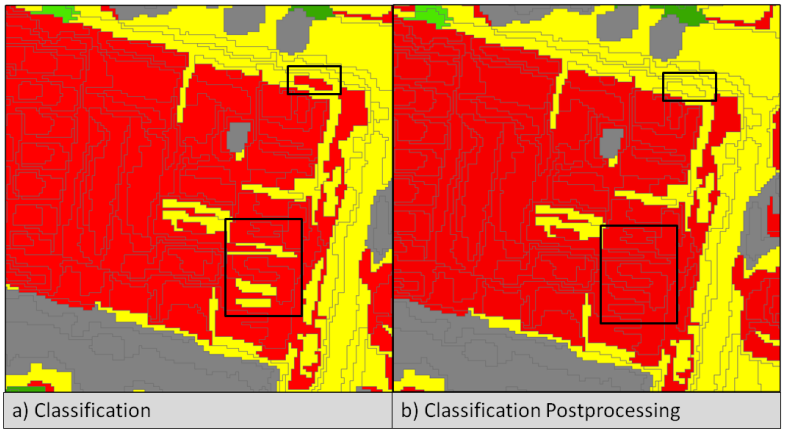


Figure 4.11.: Classification Postprocessing: Introduction of context- and spatial-based rules for confused classes of interest. a) Classification b) Classification enhanced by postprocessing steps. Red= Buildings, Yellow= Roads. Black boxes: Exemplary building and road segments that are completely surrounded by another class are assigned to that class, respectively.

The introduced rules are based on the assumption that an IO assigned to a class surrounded completely by image objects of another class most probable belongs to the class it is nested inside. Real world objects consist of several IOs (Fig. 4.6) which reinforces the supposition that an IO completely surrounded by IOs of a distinct class tends to be misclassified.

Therefore, single road segments that are surrounded by building segments or single building segments surrounded by road segments (e.g., cars) are classified as such. The same rule is applied to misclassified water segments and bare soil segments surrounded by road

or building segments, respectively.
Basically, the rules aim at assigning wrongly labeled IOs to their appropriate class, based on their spatial relationship with adjacent IOs (i.e., inclusion).

4.3. Exposure Estimation

4.3.1. Number of Buildings

For the purpose of estimating the number of buildings, an extended segmentation approach based on edge detection and morphological filtering is introduced in order to derive single buildings. The segmentation is combined with the prior classified buildings (Sec. 4.2.4) by applying a majority vote. Training samples are generated by the computation of landscape metrics and the number of automatically counted OSM-buildings, both resolved to grid cells. By the use of several regression techniques, the number of buildings for each grid cell is estimated for the whole study area and validated based on a manually generated dataset (Fig. 4.12). In the following section the individual operations will be outlined in detail.

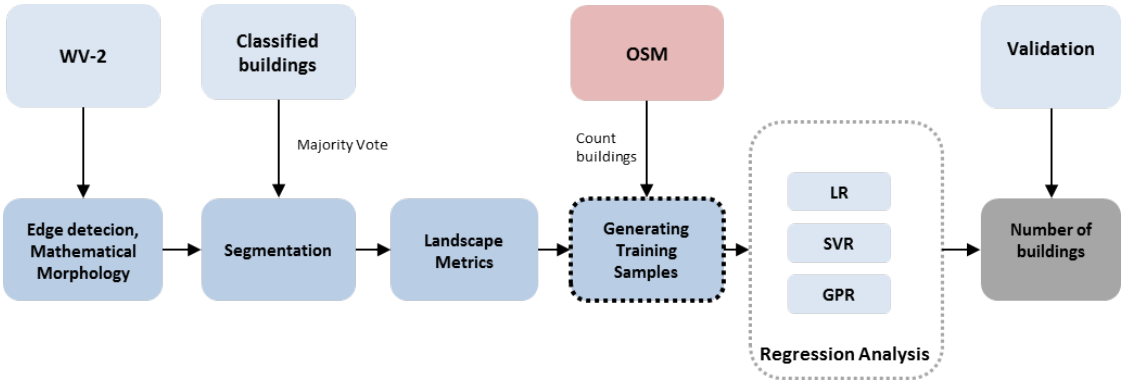


Figure 4.12.: Workflow for the estimation of number of buildings: Buildings are detected based on a new segmentation approach concerning edge detection and morphological filtering and labeled with the help of the prior building classification. Training samples are generated by computing landscape metrics for the segmentation image and the number of buildings counted in OSM. Regression analysis is conducted with Linear Regression (LR), Support Vector Regression (SVR) and Gaussian Processes for Regression (GPR) in order to estimate the number of buildings. A reference dataset is generated by the manual count of buildings.

Preparing Dataset for Building Count

Since buildings have diverse characteristics and varying appearance in VHR imagery (illumination, viewing angle, etc.), it is difficult to detect separate buildings from the data. Sirmacek and Unsalan [2011] and Wei et al. [2008] point out the major difficulties that arise:

- Buildings may be imaged from different viewpoints
- They may not have a unique representation
- They may have complex interaction with the environment and each other (e.g., shade)
- Illumination and contrast in the image may not be sufficient to detect them reliably
- Buildings do not have standard size and shape

To overcome these difficulties, an extended segmentation approach is introduced in order to derive single buildings as accurate as possible for the purpose of estimating their amount.
It is aimed at determining IOs that correspond to the size and shape of the buildings in the urban area. In addition to the multispectral channels, information from edge detection and morphological filtering is considered in the segmentation process. Similar approaches have already successfully been used in recent studies [Epifanio and Soille 2007, Pina and Barata 2003, Ali and Clausi 2001, Benediktsson et al. 2003, Kaur and Garg 2011, Peijun et al. 2007, Pesaresi and Benediktsson 2001, Sheng et al. 2009].
Thus, simple erosion, dilation, opening and closing operators with a SE size of 5 pixels (also see Sec. 4.2.2) are considered as well as canny and lee sigma edge detection for the pan channel.
Lim [1990] defines an edge in an image as a boundary or contour at which a significant change occurs in some physical aspect of the image. By means of edge detection methods, these local intensity values are tried to be detected [Ali and Clausi 2001]. The canny edge detector proposed by Canny [1986] is based on computing the squared gradient magnitude. Local maxima of the gradient magnitude that are above a certain threshold are identified as edges. Thereby, pixels with higher values than a certain threshold (non-edge pixels) are first detected and then marked again as edge node pixels if they exceed a lower threshold. Afterwards, edge detection values lower than the second threshold are removed excluding low intensity gradient edges from the result [Ali and Clausi 2001, Trimble GmbH 2014]. In contrast, the edge Extraction Lee Sigma algorithm creates two layers using a specific edge filter extracting bright and dark edges. Thereby, the sigma value describes how far a data point is from its mean in standard deviations. A high sigma value results in a more sensitive edge detection [Trimble GmbH 2014].
Finally, a majority vote is applied in order to transfer the prior building classification to the new segmentation image.

Landscape Structure Metrics

For the determination of the number of buildings several characteristics were derived making use of landscape structure metrics (LSM). Structural relationships can be expressed

by indices characterizing the shape, pattern and arrangement of landscape elements, spatially depicted as individual IOs [Berry 2007].
Therefore, the study site is divided by a grid of 250 m height/length according to administrative census boundaries of the community Valparaíso. For all segments having their centroid in a given grid cell, statistics and metrics are calculated and assigned to the grid. In this way, double counting of segments (segment areas falling in more than one grid cell) is prevented. Table 4.4 depicts the derived metrics for this analysis. Area metrics help to describe fragmentation and variation of the segments and their area. They deal with the size and amount of the segments within one grid cell [Berry 2007]. The quantification of the amount of segments includes the number of segments per grid cell, the percental coverage of built-up area per grid cell and the amount of segments per grid cell relative to the whole study site. With respect to the size of segments, the average size of the segments within a grid, the largest and smallest segment as well as the range of segment sizes per grid cell are quantified. The standard deviation and the coefficient of variation of the segment size depict the degree of spread around the mean segment size and the variation of segment sizes within a grid cell, respectively. On the one hand, edge metrics quantify the segment boundaries within a grid cell while, on the other hand, shape metrics summarize boundary configuration, measuring the complexity of segments boundaries [Berry 2007]. The chosen shape index measures the complexity of a segment compared to a reference shape (square). An increasing shape index signifies an increasing difference from square shape and, thus, more irregular structured segments. Finally, density measures consider the frequency of occurrence and serve as first order indicator of spatial heterogeneity (in this case buildings and others). Overall, each IO will be represented by a 12-dimensional feature vector.

Multivariate Regression Analysis

The problem of estimating the number of buildings from the remote sensed imagery is viewed as a regression issue where it is aimed at reproducing the relationship between a set of extracted features and the number of buildings. According to limited available OSM data, 25 training cells are selected for the analysis and buildings are counted automatically. For the generation of an independent reference dataset, 25 grid cells are randomly selected and buildings are counted manually. Both, training and validation datasets comprise the number of buildings (dependent variable) that is to be predicted for the whole dataset, and computed LSM (independent variables, Tab. 4.4). The target dataset on which the model will be applied later on contains only the LSM.
Three different regression approaches are considered to solve the regression problem. Linear Regression (LR) is a standard approach of multivariate statistics. Gaussian Pro-

	LSM	Description
Area metrics	Segment density	Number of segment per grid cell
	Mean Segment size	Average area of all segments per cell
	Segment standard deviation	Variation in segment size
	Segment coefficient of variation (%)	Variation in segment size
	Min segment size	Size of smallest segment
	Max segment size	Sizes of largest segment
	Range segment size	Range of segment sizes per grid cell
	Percent coverage	Relative area of built-up segments per grid cell
	Relative number of segments	Number of segments per grid cell relative to total number of segments in landscape
Edge metrics	Edge density	Total length of all edge segments per grid cell
Density measures	Segment density	Number of segments relative to total area per gird cell
Shape metrics	Shape index	Normalized ratio of segment perimeter to total area, adjusted by a constant for a squared standard

Table 4.4.: Landscape Structure Metrics (LSM) derived for 0.625 km^2 sized grid cells to quantify spatial patterns of building segments for Valparaíso. Computed measures can be classified in area metrics, edge metrics, density measures and shape metrics. Each IO will be represented by a 12-dimensional feature vector.

cesses for Regression (GPR) and Support Vector Regression (SVR) are non-linear approaches that have been showing promising results in recent studies [Bazi and Melgani 2010, Camps-Valls et al. 2006, Chen et al. 2010, Deswal 2011, Guo et al. 2014, Okujeni et al. 2013, Pasolli and Melgani 2011, Pasolli et al. 2008, Tuia et al. 2011, Verrelst et al. 2012]. The aim is to find possible relationships between the dependent variable "Number of buildings" and the independent variables (LSM) to predict the number of buildings for the entire city of Valparaíso per grid cell.

Linear Regression is based on the assumption of a linear coherence between a dependent variable y and one or several independent variables x among which possible relations are to be determined [Backhaus et al. 2011]. Therefore, a regression line containing the predicted values is determined. It is the optimal line crossing the observed point cloud

built by x and y.

$$\bar{y_i} = b_0 + b_1 x_i \tag{4.2}$$

The determination of the regression line requires the determination of the parameters b_0 and b_1 in a way that it optimally displays the linear tendency in the scatterplot [Bahrenberg et al. 1999]. This is achieved based on the gaussian method of ordinary least squares (OLS). Thereby, value pairs are approximated optimally with the smallest possible deviation by minimizing the sum of the quadratic deviations of the line to the observed values. According to equation (4.3) b_0 and b_1 have to be chosen so that $E = E(b_0, b_1)$ reach a minimum [Bahrenberg et al. 1999].

$$E = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2 \rightarrow \text{Min} \tag{4.3}$$

In the regression function, the given variables are multiplied with the determined constant coefficients. The combined sum of these terms is called a linear combination. A multiple linear regression considers several independent variables. The main drawback of a standard linear approach is that it only allows limited flexibility and if the relationship between input and output cannot reasonably be approximated by a linear function, the model will give poor predictions [Rasmussen and Williams 2006]. In contrast, non-parametric regressions proceed without specific information about the regression function and determine the approximation only based on available training data.
Gaussian process for regression (GPR) is used to generate predictions of function values based on training data. Here, the key assumption is that the data can be represented as a sample from a multivariate Gaussian distribution and that adjacent observations should convey information about each other [Ebden 2008, Deswal 2011]. By the calculation of expectation values and variance estimations for the underlying function, possible deviations or error rates can be determined. Consequently, it is the generalization of a gaussian distribution over a finite vector space to a function space of infinite dimension [Ebden 2008]. Just as a Gaussian distribution is fully specified by its mean and covariance matrix, a Gaussian process is specified by a mean and a covariance function or kernel function, respectively. A GP is defined as a collection of random variables, any finite number which has a joint multivariate Gaussian distribution [Deswal 2011]. Let $\chi * \gamma$ represent the domain of inputs, from which n pairs (x_i, y_i) are drawn independently and identically distributed. Then, for regression it is assumed that

$$y \subseteq \Re \tag{4.4}$$

A GP on χ is further defined by a mean function

$$\mu : \chi \rightarrow \Re \tag{4.5}$$

and covariance function

$$k : \chi * \chi \rightarrow \Re \tag{4.6}$$

The main assumption of GPR is that y is defined by $y = f(x) + \xi$ where $\xi \sim N(0, \sigma^2)$ [Deswal 2011]. ξ is the normal independent and identically distributed observational error. Each point in the function space $f(x)$ is treated as a random variable and the covariance function shows the statistical dependency between the function values. Predictions produced by a Gaussian process depend entirely on the covariance matrix, in other words, the problem of learning in a Gaussian process is the problem of finding suitable properties of the covariance matrix [Mackay 1998]. Therefore, when training a GPR model, a suitable covariance function and its parameters has to be chosen. Detailed theoretical background of Gaussian Processes is given in Ebden [2008], Mackay [1998] and Rasmussen and Williams [2006].

Support Vector Machines (SVM) are also supervised, non-parametric statistical learning techniques designed to solve classification and regression problems [Smola and Schölkopf 2004]. The main idea of support vector regression (SVR) is to estimate the linear dependency between pairs of n-dimensional input vectors and 1-dimensional target variables by fitting an optimal approximating hyperplane to a set of training samples. The hyperplane is defined by a linear function that is found based on structural risk minimization. Thereby, a specified derivation threshold ε must not be exceeded. If it is exceeded, training instances become support vectors and will be linearly penalized by the degree of distance to ε [Smola and Schölkopf 2004]. Therefore, the goal is to find a function $f(x)$ that has at most ε deviation from the actually obtained targets for all training data and at the same time is as flat as possible, which is seeked by a small ω [Smola and Schölkopf 2004]. However, sometimes optimization is not feasable and therefore slack variables ξ_i, ξ^* are introduced to allow for some errors, avoid overfitting and to cope with otherwise infeasible constraints of the optimization problem. C is a regularization variable for the trade-off between the function flatness and the tolerated error of incorrectly estimated sample instances (amount up to which deviations larger than ε are tolerated). The optimal parameters for C and ε have a decisive influence on the generalization performance and the model complexity and have to be optimized empirically. According to Vapnik [1995] the resulting minimization objective can be described as follows:

$$\text{minimize} \qquad \frac{1}{2}||w|| + \sum_{i=1}^{l}(\xi_i + \xi_i^*) \qquad (4.7)$$

$$\text{subject to} \qquad \begin{cases} y_i - \langle w, x_i \rangle - b & \leq & \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i & \leq & \varepsilon + \xi_i^* \\ \xi, \xi^* & \geq & 0 \end{cases} \qquad (4.8)$$

Thereby, the linear model is optimal when it minimizes a cost function considering a maximized margin that encloses the samples, and a minimized error of approximation [Okujeni et al. 2013]. Since it turns out that in most cases the optimization problem can be solved more easily in its dual formulation, a dual set of Langrange variables is introduced. The dual optimization problem is then solved by quadratic programming techniques and the final estimation function or support vector expansion is given by

$$f(x_*) = \sum_{i=1}^{n}(a_i - a_i^*)K(x_{i'}, x_*) + b \qquad (4.9)$$

where a_i a_i^* refer to the remaining Lagrange multipliers, $x_{i'}$ and x_* to the training data, n to the number of support vectors, b to the slope and K to the the Kernel-function [Smola and Schölkopf 2004]. Further information about SVR are provided in Drucker et al. [1996], Smola and Schölkopf [2004], Vapnik [1995], Verrelst et al. [2012].

In the experimental setup, LR is performed accomplishing an attribute selection using the M5 method and eliminating collinear attributes at the same time. The M5 method builds trees whose leaves are associated to multivariate linear models and removes the attributes with the smallest standardized coefficient until no improvement is observed in the estimate of the error given by the Akaike information criterion [Wang and Witten 1997]. Multicollinearity means that a regressor must not constitute a linear function of the other regressors. If that is the case, the regression analysis is not computational feasible and interpretable [Backhaus et al. 2011].
As discussed above, a necessary hyper-parameter optimization for GPR and SVR was performed deploying a GridSearch in WEKA using RMSE as measure to be minimized. For SVR the sequential minimal optimization (SMO) algorithm for support vector machines (regSMOImproved) is used which normalizes the training data by default [Smola and Schölkopf 2004]. As optimal complexity parameter, C = 2 is determined. A polynomial Kernel is chosen with an optimal exponent $\varepsilon = 1$.
Hyper-parameter optimization for GPR is concerned with modeling covariance. Thereby, we searched for best performing parameters for the kernel and the noise parameter for the GP. GPR also normalizes data by default. A RBFKernel function is found appropriate

with $\gamma = 0.4$. Attribute selection did not increase results for both, therefore all variables were considered.
LR and SVR require a post-processing step. Both methods partly output unrealistic negative or super-positive prediction values. The analysis of the spatial distribution of unrealistic predictions reveals that all of them are associated with grid cells of poor segment coverage (up to only one segment per cell) at borders of urban agglomerations. Further analysis reveal that they comprise misclassifications. Given this rather meaningful location and characteristics of outliers, these values can be excluded from the result to achieve physically meaningful values for building number estimations.

4.3.2. Spatial Disaggregation and Estimation of Population

Natural hazards are highly variable over time and space and their effects on the building stock and population can be on small or large scales, depending on magnitude of the disaster and vulnerability of the assets (see Sec.2.1). The demand for census data within the scope of exposure estimation for disasters is independent of administrative units [Steinnocher et al. 2011]. Therefore, only fine scale population datasets can provide an accurate estimate of the population actually exposed to a risk.
Subsequently, areal interpolation methods are required to transform census population data from a coarse spatial unit (source area) to a smaller unit (target area). By means of disaggregation techniques a refined population surface is obtained [Wu et al. 2005].

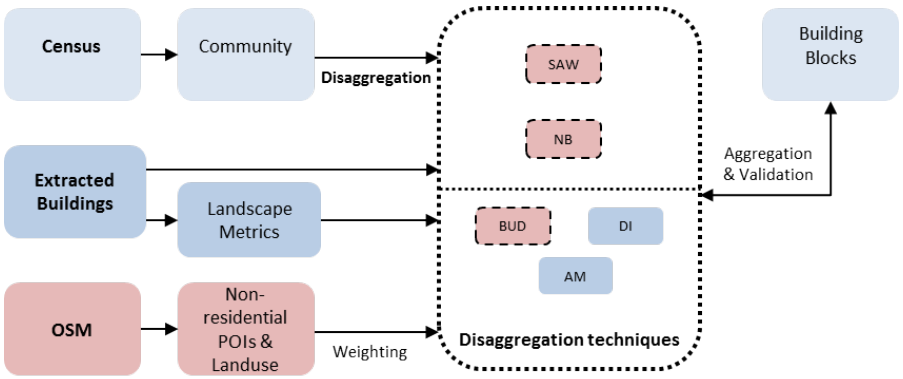


Figure 4.13.: Workflow for spatial disaggregation and estimation of population: Census data on community level are distributed to pixel level by means of different disaggregation algorithms. Population distribution is enhanced by ancillary information from the remotely sensed data and OSM. Validation is conducted by means of census data on building block level.

The generated land cover classification serves as basis for the disaggregation and estimation of population in Valparaíso. All tested algorithms are based on the assumption that people only dwell in areas classified as buildings. Thus, the sum of the extracted buildings depict the source area. The tested disaggregation algorithms can be divided in

methods without ancillary information and methods making use of ancillary information, specifically landscape metrics (Fig. 4.13). Moreover, OSM data is used in order to enhance disaggregation by weighting areas depending on their function or usage type.
Census data on community level for 2002 constitute the data basis. Census data from the same year on building block level serve for validation. Building blocks falling inside the burnt area are excluded from the whole analysis (Sec. 3.2.2 and 4.1.2). Population numbers in this area are subtracted from all census levels. Furthermore there is no population information on building block level available for the district Cerro las cañas (Fig. 3.5) and it is therefore excluded for validation purposes. The validation is accomplished on the base of the remaining 2906 building blocks.

Disaggregation without Ancillary Information

In a first step a simple linear population distribution or simple areal weighting (SAW) is conducted [Wu et al. 2005]. Areal weighting is based on the assumption that population is uniformly distributed within the source zone. Therefore population is assigned to each built-up segment proportional to its size [Wu et al. 2005]. In equation (4.10), A_i is the area of an image pixel or image object respectively, $\sum A_i$ is the sum of all areas and Pop the population on community level.

$$Pop_i = \frac{A_i}{\sum A_i} * Pop \tag{4.10}$$

Thus, the estimation is based on three assumptions:

1. The population is distributed equally over the entire region
2. People only dwell in buildings classified as such
3. Local disaggregation of population is dependent on the interrelationship of building area and population

Disaggregation with Ancillary Information

Population can also be related to other information that can be used to assist disaggregation [Wu et al. 2005]. The distribution of population following structural characteristics of an urban morphology is determined by assuming a dependency between numerical and spatial units. The idea of so-called dasymetric methods is to use knowledge of the locality to identify areas with zones of different population densities [Mennis and Hultgren 2006, Wu et al. 2005]. Dasymetric zones are generated by using ancillary information that can be extracted from the remote sensing data [Eicher and Brewer 2001].
Subsequently, as a first approach the beforehand estimated number of buildings is utilized

(see Sec. 4.3.1). Building number estimations are available based on a 250 × 250 m^2 grid. By dividing the overall population by the estimated number of buildings, the average number of inhabitants per house can be derived. By means of this information the population can be calculated by multiplying the average number of persons per buildings with the number of buildings per grid cell. In a second step, the resulting number of population per grid cell is disaggregated to pixel level by SAW. The idea behind is to improve the SAW by controlling the population distribution beforehand based on ancillary information.

As a second approach, the method for population disaggregation suggested by Steinnocher et al. [2011] is conducted. It is a dasymetric mapping approach that utilizes the built-up density as ancillary data to quantify homogenous spatial units. The approach is based on the following assumptions [Steinnocher et al. 2011]:

1. The population density is proportional to the built-up density
2. People only dwell in buildings classified as such
3. within a region the ratio of population- and built-up density is constant

Therefore, the built-up density is calculated for the units of building blocks and classified into three classes, where density values under 40% are assumed as low built-up density, values between 40% and 60% as medium density and values greater than 60% as high density class.
Assuming that the built-up density influences the population density in a proportional manner, their relationship can be represented as follows, where PD is the population density, BUD the built-up density and k the weighting parameter.

$$PD = k * BUD \tag{4.11}$$

The total population in a spatial unit Pop_i is then calculated using the area and the weighted density (Eq. 4.12). The prevenient calculation of the weighting parameter k therefore includes the built-up density and the area per density class, respectively.

$$Pop_i = \sum_i A_i * k * BUD_i \tag{4.12}$$

Furthermore, additional parameters for the equation of disaggregation based on built-up density are tested. Ancillary information, that can be further extracted from the previous classification results, is utilized. The idea is to benefit from landscape metrics that can be calculated over a building block unit in order to describe the settlement dispersion (see Sec. 4.3.1). The assumption of the approach is that the population density behaves

inversely proportionate to a settlement dispersion.
Initially, the population is distributed taking into account the Dispersion Index (DI) according to Wiesner and Taubenböck [2014]. The model allows measuring spatial compactness or spatial dispersion of two-dimensional settlement patterns with two parameters: the number of segments (NS) and the largest patch in a spatial unit (LS) (Eq. 4.13 and 4.14).

$$NS_i = \sum S_i \tag{4.13}$$

$$LS_i = \frac{\max A_i}{\sum\limits_{i=n}^{n} A_i} * 100 \tag{4.14}$$

Both parameters are normalized per unit area to NS_{norm} and LS_{norm} in order to facilitate the quantification of the degree of dispersion. Finally, the DI is a combination of the two spatial metrics.

$$DI_i = \frac{NS_{norm} + (100 - LS_{norm})}{2} \tag{4.15}$$

To obtain more significant indicators for the population distribution, the DI is further standardized setting the minimum and maximum border of standardization to the smallest and biggest DI values, respectively. A new population disaggregation result is obtained by combining the equation of the built-up density (Eq. 4.12) with the influencing factor of the DI. It is amplified by 0.5 in order to obtain values between 0.5 and 1.5.

$$Pop_i = (\sum_i A_i * k * BUD_i) * (DI + 0.5) \tag{4.16}$$

Finally, the area weighted mean (AM) is calculated over the segments per building block. The AM summarizes and averages the segments in one BB weighting their size (Equation 4.17). Subsequently, large segments have a higher influence than small segments. In contrast to the DI, the AM is not converse to the population density, a high AM signifies a small dispersion and consequently a high population density. The equation for built-up density is then extended by the influencing factor for the AM, just like the DI (Eq. 4.18).

$$AM = \sum A_i * \frac{A_i}{\sum A_i} \tag{4.17}$$

$$Pop_i = (\sum_i A_i * k * BUD_i) * (AM + 0.5) \tag{4.18}$$

Disaggregation with OSM Data as Ancillary Information

Finally, the population distribution is aimed to be improved by integrating OSM-data. Crooks et al. [2015] and Kunze and Hecht [2015] have recently emphasized the role of crowdsourcing in the study of urban morphology. For this purpose, landuse object and Points of Interest (POIs) other than residential are extracted. OSM landuse objects are polygons implying the type of landuse and POIs can be associated with nodes, ways or polygons, especially indicating public-, recreational-, and cultural places.
The implementation of OSM data in the population disaggregation is based on the assumption that different categories of landuse objects and POIs indicate distinct usage types of buildings and therefore hint at population distributions. Furthermore, OSM landuse objects can be used to identify misclassified areas that are not actually buildings (e.g., parking areas with cars or cemeteries). Therefore, additional information of OSM objects were classified in the predefined meaningful categories industrial, culture, leisure, gastronomy, education, traffic, health, tourism, services, shopping, religion and public (see Appendix A.2, Tab. A.1). Only objects with usage types other than residential were selected. The OSM information is then intersected with the previously extracted building objects. For point objects the intersection is conducted within a distance of 3 meters (Fig. 4.14).
Concerning the OSM information and its relation to population statistics, different assumptions can be made:

1. Areas covered by developed OSM categories are unpopulated
2. Some OSM categories refer to unpopulated areas and some to less populated areas.

For assumption 1), particular areas are weighted by the factor 0 assuming that no people live in these areas. For assumption 2) the categories gastronomy, services, shopping, tourism and health are assumed to be medium populated for the reason of either temporary residents or the hypothesis of partial use, like e.g. commercial usage of the main floor and residential use of further storeys. Thereby, the categories are weighted by the factor 0.5, respectively. Remaining categories are weighted with 0 as above.
OSM ancillary information serve as additional parameters for population distribution and are therefore combined with the equations for BUD, SAW and BUD (Equations 4.10 and

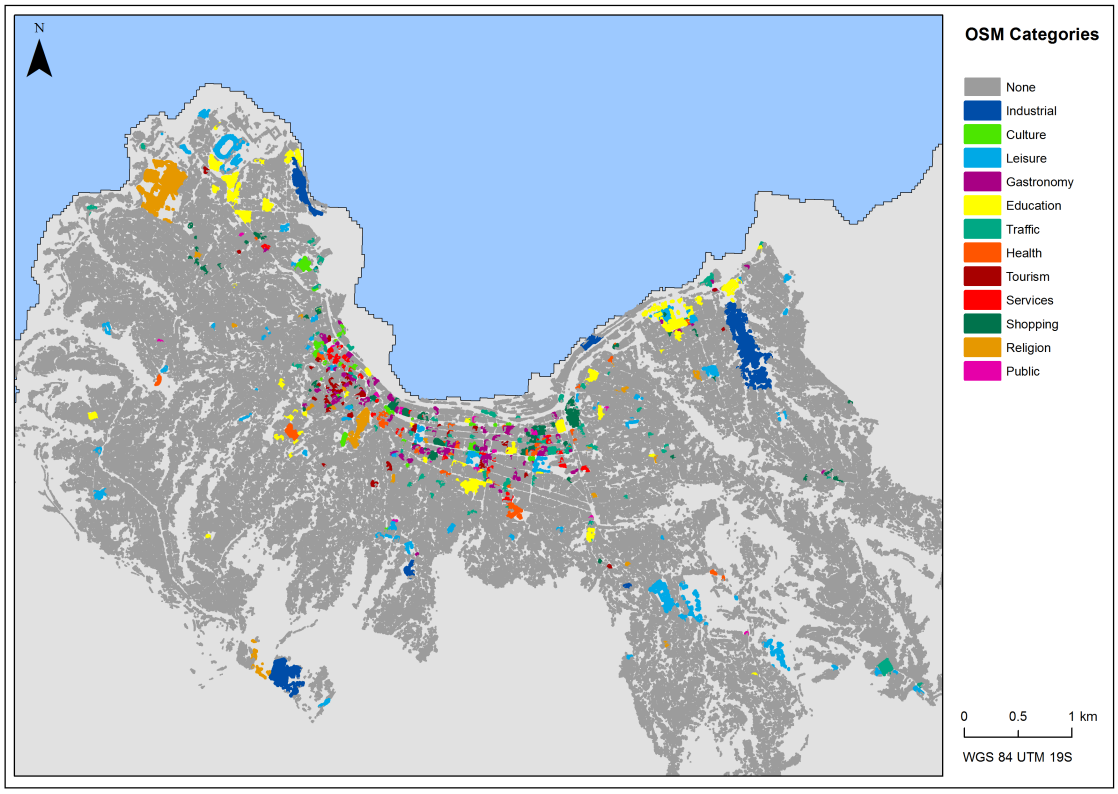


Figure 4.14.: OSM categories: OSM landuse information and POIs are classified in 12 predefined categories with respect to their type of usage.

4.12), considering merely segments with assigned OSM categories.

4.4. Measures for Accuracy Assessment

It is widely accepted that no classification or regression analysis is complete until its accuracy has been rigorously assessed [Foody 2004]. In the following chapter, accuracy assessment measures and techniques are elucidated for both, information extraction and exposure estimation.

4.4.1. Assessment of Information Extraction Accuracy

In order to assess the performance of the classification model with regard to the different classes, the confusion matrix (CM) is calculated. A CM comprises all sample units assigned to a particular class in the classification relative to the number of samples assigned to that class in reality. Usually, the columns represent the reference data while the rows indicate the classification generated from remote sensed data. The sum of the

columns therefore represents the number of samples per class in the reference data while the row sum represents the number of samples assigned to each class in the classification process. The CM is an effective way to represent map accuracy describing the individual accuracies of each class along with commission errors (error of inclusion) and omission errors (errors of exclusion) [Congalton and Green 1999]. A commission error is defined as including an area into a class where it does not belong to. In contrast, the ommission error is excluding an area where it truly belongs, i.e. the classification technique has failed to classify them in the correct class [Congalton and Green 1999]. Based on the CM different accuracy measures can be calculated.

The Overall Accuracy (OA) approximates the accuracy of the entire classification in one single value by summing up the correctly classified sample units (CM diagonal) divided by the total number of samples in the matrix. The calculation is represented in equation (4.19) where n is the number of samples and n_{ij} the number of samples classified into class i in the remotely sensed class and class j in the reference dataset.

$$OA = \frac{\sum_{i=1}^{n} n_{ij}}{n} \tag{4.19}$$

Thereby, a sample of locations on the map is compared with the same locations on the reference data keeping track of the number of times there was an agreement [Congalton and Green 1999]. In this context, according to Congalton and Green [1999], an OA above 85% can be considered as an acceptable result.

Moreover, Producer's and User's Accuracy are calculated in order to represent individual class accuracies. They are defined as follows in equations (4.20) and (4.21), where n_{+j} are the number of samples classified into class i in the remotely sensed classification and n_{i+} the number of samples classified into class j in the reference dataset.

$$PA_j = \frac{n_{jj}}{n_{+j}} \tag{4.20}$$

$$UA_i = \frac{n_{ii}}{n_{i+}} \tag{4.21}$$

PA and UA should always be examined together. PA refers to the probability that a certain land-cover of an area on the ground is classified as such, while the UA refers to the probability that a segment labeled as a certain land-cover class belongs to this class in reality.

Finally, the most common measure for global accuracy of classifications is the κ-statistics [Foody 2002]. The Kappa analysis is a standard component of nearly every accuracy assessment in remote sensing and statistically explains how well two thematic datasets

Kappa κ	Agreement
< 0	Less than chance agreement
$0.01 - 0.20$	Slight agreement
$0.21 - 0.40$	Fair agreement
$0.41 - 0.60$	Moderate agreement
$0.61 - 0.80$	Substantial agreement
$0.81 - 0.99$	Almost perfect agreement

Table 4.5.: Interpretation of κ statistics [Landis and Koch 1977]

match [Congalton and Green 1999]. κ is defined by equation (4.22) where n is the sum of all objects in the confusion matrix (number of references), r the number of rows of the CM (number of classes), x_{ii} the quantity of correctly classified objects (CM diagonal), x_{i+} the sum of all objects per row and x_{+i} the sum of all objects from a column [Foody 2002].

$$\kappa = \frac{n \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} x_i * x_{+1}}{n^2 - \sum_{i=1}^{r} x_i * x_{+i}} \tag{4.22}$$

κ considers commission and omission errors and is therefore not biased concerning the distribution of the samples [Foody 2002].
The κ index varies between 0 and 100%, where 0% indicates that the level of agreement is equal to the agreement due to chance and 100% indicates perfect agreement. Table 4.5 outlines the interpretation of resulting κ statistics according to Landis and Koch [1977].

4.4.2. Assessment of Exposure Estimation Accuracy

In the following, selected measures for the quality assessment with respect to the exposure estimation are illustrated. For the assessment of regression analysis the most known measure is the correlation coefficient. In statistics, the correlation coefficient r_{xy} measures the statistical correlations between the predicted values and the actual values [Bahrenberg et al. 1999]. It can result in values between -1 and +1, whereby a value of 0 indicates no correlation. The closer $|r_{xy}|$ to 1 the stronger the linear correlation [Bahrenberg et al. 1999]. For two variables x and y it is defined as

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x}) * (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 * \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{4.23}$$

The mean absolute error (MAE) quantifies the adaption of the regression function to the empirical data (so called "goodness of fit") by measuring the average magnitude of errors in a set of forecasts, without considering their direction [Backhaus et al. 2011, Makridakis and Hibon 1995]. It is a linear score, which means that all the individual differences are weighted equally in the average. Here, x_i constitutes the actual value, $\hat{x}_i$ the estimated value and n the number of values, i.e. the number of building blocks. This definition of variables will be consistent for further error measures presented in this section.

$$\text{MAE} = \sum_{i=n}^{n} \frac{|x_i - \hat{x}_i|}{n} \tag{4.24}$$

In contrast, the root mean squared error (RMSE) is a quadratic scoring rule which measures the average magnitude of the error. The difference between forecast and corresponding observed values are each squared and then averaged over the sample taking the square root. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. The RMSE varies with the variability of the error magnitudes and the total-error or average-error magnitude (MAE), and the sample size n [Chai and Draxler 2014]. Thus, the RMSE will increase as the variance associated with the frequency distribution of error magnitude increases [Res et al. 2005]. This means, the RMSE is most useful when large errors are particularly less desirable than two or more smaller ones whose sum is about the same as the large error [Makridakis and Hibon 1995]. Thus, in contrast to the MAE, the RMSE is very sensitive to outliers (large errors).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=n}^{n}(x_i - \hat{x}_i)^2}{n}} \tag{4.25}$$

The MAE and the RMSE can be used together to diagnose the variation in the errors in a set of forecasts. The RMSE is by definition always larger than the MAE; the greater difference between them, the greater the variance in the individual errors in the sample [Chai and Draxler 2014]. The ratio of the RMSE and MAE can range and an inconsistency between MAE and RMSE can also appear [Res et al. 2005].

Moreover, the MAPE is a relative measure that expresses errors as a percentage of the actual data [Makridakis and Hibon 1995]. It is a good measure to compare different methods since the MAPE of each tells about the average relative size of their errors. Its disadvantage is its sensitivity to scale, which makes it unfavorable or even worthless

when working with low-volume data [Makridakis and Hibon 1995]. Since the observed value x_i is in the denominator of the equation, it is undefined when this value is zero, and will take on extreme values when x_i is quite small. These low values can dominate the statistic which is undesirable.

$$\text{MAPE} = \frac{1}{n}\sum_{i=n}^{n}\frac{|x_i - \hat{x}_i|}{x_i} \tag{4.26}$$

The symmetric mean absolute percentage error (SMAPE) improves the MAPE by correcting its problem of asymmetry and strong influence of outliers [Makridakis and Hibon 1995]. This is achieved by dividing the forecasting error by the average of both. The SMAPE has a lower and an upper bound: Equation (4.27) provides a result between 0 and 200 %. Since a percentage value between 0 and 100 % is easier to interpret the formula for SMAPE in Equation (4.28) is often used [Makridakis and Hibon 1995].

$$\text{SMAPE}_1 = \frac{1}{n}\sum_{i=n}^{n}\frac{|x_i - \hat{x}_i|}{(x_i + \hat{x}_i)/2} \tag{4.27}$$

$$\text{SMAPE}_2 = \frac{1}{n}\sum_{i=n}^{n}\frac{|x_i - \hat{x}_i|}{x_i + \hat{x}_i} \tag{4.28}$$

By means of the root relative squared error (RRSE) and the relative absolute error (RAE), it can be assessed how much a scheme improves on simply predicting the average.
The RRSE is computed by dividing the RMSE by the RMSE obtained by just predicting the mean of target values (and then multiplying by 100). Therefore, smaller values are better and values $> 100\%$ indicate a scheme is doing worse than just predicting the mean.

$$\text{RRSE} = \sqrt{\frac{\sum_{i=n}(\hat{x}_i - x_i)^2}{\sum_{i=n}(x_i - \bar{x}_i)^2}} \tag{4.29}$$

The RAE is a measure of uncertainty of a measurement compared to the size of the measurement. It is computed by normalizing the total absolute error by the total absolute error of the simple predictor (Eq. 4.30). Thus, values of $|\hat{x}_i - x_i|$ or $(\hat{x}_i - x_i)^2$ reveal how much x differs from its mean value and therefore giving a result related to the scale of x. Further information on how errors measures are computed in WEKA can be obtained in WEKA Data Mining [2015].

$$\text{RAE} = \frac{\sum_{i=n}|\hat{x}_i - x_i|}{\sum_{i=n}|x_i - \bar{x}_i|} \tag{4.30}$$

5. Results & Discussion

In the following chapter, results of previous analysis are revealed and discussed. Primarily, validation efforts of the classification of satellite imagery based on OSM-samples are revealed and analyzed (Sec. 5.1.1). The conducted exposure estimations are discussed in the subsequent section (Sec. 5.2). Finally, the developed exposure dataset is applied within a tsunami scenario (Sec. 5.3).

5.1. Information Extraction

5.1.1. Classification

For the generation of a LULC map of the study site, a RTF classifier model was trained based on OSM-samples and applied to the entire remote sensing imagery (Sec. 4.2.4). Based on the model performance, classification accuracies for the test areas can be estimated. Thereby, an overall accuracy (OA) of 85.9% and a κ of 80.8% are achieved. Such accuracy levels can be considered as acceptable and substantial result according to Congalton and Green [1999] and Landis and Koch [1977].
The confusion matrix was generated from manually gathered labeled samples (Sec. 4.2.4). It allows for further examination of class-specific accuracies. The classes of great interest with respect to further analysis and application considerations are the built-up classes "roads" and "buildings", which are generally difficult to delineate (both are man-made objects with similar characteristics in complex urban structures). 13.5% of building reference

LC Class	**Buildings**	**Roads**	**Trees/ Bushes**	**Meadow**	**Water**	**Bare Soil**	**Sum**
Buildings	283	23	3	0	0	28	337
Roads	42	65	0	1	0	15	123
Trees/Bushes	3	3	389	3	0	4	402
Meadow	6	0	4	65	0	6	81
Water	0	0	3	0	35	1	39
Bare soil	0	3	0	0	0	71	74
Sum	334	94	399	69	35	125	1056

Table 5.1.: Confusion matrix of classification: 3600 training samples and 1056 reference instances. Diagonal highlighted in color: Correctly classified instances.

	PA	UA
Buildings	84.7%	83.9%
Roads	69.1%	53.8%
Trees/Bushes	97.4%	96.7%
Meadow	94.2%	80.2%
Water	100%	89.7%
Bare soil	56.8%	95.9%

Table 5.2.: Producer's and User's accuracies of classification

instances and 30.9% of road instances are assigned mistakenly to other classes, in particular to the classes "roads" or "buildings", respectively. While the classes "Trees/bushes", "meadow" and "water" show reliable accuracies, the class "bare soil" features non-negligible confusion with all other classes (43.2%), especially "roads" and "buildings" (Fig. 5.1).
Table 5.2 illustrates the Producer's and User's Accuracy (PA, PU) for the classification. As already explained in section 4.4.2, the PA refers to the probability that a certain LULC class of an area on the ground is classified as such, while the UA refers to the probability that an instance labeled as a certain LULC class belongs to this class in reality. Except for the class "bare soil", the PA's show higher values than the UA's. This features that there are in general more omission errors than commission errors. In other words, areas on the ground are rather excluded from the class they really belong to than included into a class they do not belong to. Classes with most confusion are "buildings", "roads" and "bare soil". "Bare soil" has a relatively high commission error, especially with respect to "buildings" and "roads".
84.7% of the buildings and 69.1% of the roads on the ground are classified as such, while 83.9% of classified buildings and only 53.8% of classified roads actually belong to the class they have been assigned to. This signifies especially for the class "roads", that more objects are classified as "roads" than actually are roads on the ground. This is described by its commission error of 30.9% that is dominated by confusions with "buildings". In contrast, the class "buildings" only has a commission error of 13.5%. Yet, for both, the omission error is higher than the commission error with 16.1% for buildings and 47.2% for roads. "Bare soil" is the class of highest confusion with respect to "roads" and "buildings". Its rather high commission error of 43.2% and at the same time low omission error of 4.1%, highlights the inclusion of false classes (mainly "buildings" and "roads") rather than excluding bare soil instances from its own class. Therefore, it becomes obvious that the classes of main interest, i.e. "roads" and "buildings", are frequently confused with "bare soil" while vice versa, bare soil is rarely classified as "roads" or "buildings". This problem probably occurs when both, bare soil and buildings, reveal very high reflectance values. Anyway, it has been noticed in previous studies, that the heterogeneity of sealed surfaces

causes numerous confusion with bare soil [Thomas et al. 2003]. Therefore, it is suggested to consider bare soil as part of the sealed surface classes (arguing with the high degree of imperviousness of bare soil in urban areas) [Weng and Hu 2008] or to use hyperspectral data [Herold et al. 2002a, Misakova et al. 2006, Syed et al. 2005]. Also dark building materials, shadow and water are often difficult to distinguish [Dare 2005, Van de Voorde et al. 2007].

Sensitivity of the Classification Model to the Number of Training Samples

In order to assess the sensitivity of the RTF classification method to the number of training samples, five experimental scenarios were set up for the dataset: For each LULC class, the sample size was reduced to 300, 150, 75, 38 and 18. For each test scenario, samples are drawn randomly out of the 600 possible samples per class, while keeping the initial dataset unchanged. In order to reevaluate the model with respect to different training sizes, the experiments were repeated 10 times each and the respective results per scenario were thus averaged (Fig. 5.1).

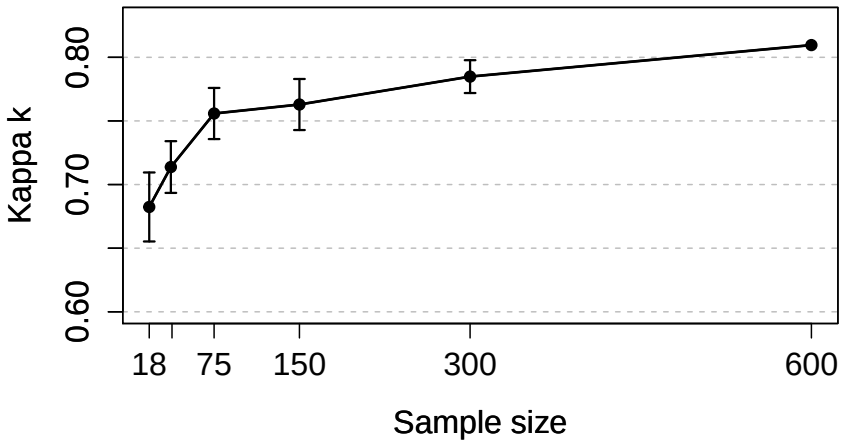


Figure 5.1.: Sensitivity of RTF method to the number of training samples: Functions of κ for different sizes of the training set. The mean κ for each experiment is illustrated as well as the standard deviation (error bar).

As expected, the κ exhibits a decreasing behavior for a decreasing training set size (Fig. 5.1). The κ increases monotonously with increasing training set sizes until a κ of 0.8 is reached. 18 to 75 training samples per class induce a strong increase of κ from 0.68 to 0.76. It exhibits a plateau effect for training sets greater than 75. This features a decreasing improvement when the training set has reached a certain size. With more than

75 samples, the achieved improvement decreases with the increase of additional samples. This effect is accompanied by an increasing standard deviation with a decreasing training set size. With a training set size of 300, the standard deviation of κ is notably smaller than the one of smaller training set sizes. This effect features greater stability and less variation with increasing training set sizes.
In summary, the κ-function highlights increasing reliability and quality of the classification model with growing training set sizes. Thereby, 600 training samples represent an acceptable quantity for the analysis, since it cannot be assumed that larger training sets will show significantly better results (plateau effect). Yet, the model performance is sensitive to particular small training set sizes which is an important finding, especially when analyzing areas with limited OSM availability.

5.1.2. Classification Postprocessing

As stated in section 4.2.5, classification postprocessing intends to enhance classification results by particularly focusing on confused classes and classes that are of interest with respect to further analysis. In section 5.1.1, resulting classification accuracies were examined. Subsequently, this chapter deals with examining the enhancement of classification accuracies achieved by postprocessing, in particular for the classes buildings and roads.
Overall, by means of the post-classification step, an improvement of 0.8 % could be achieved for the OA, resulting in an OA of 86.7%. The κ-statistics could be improved by 1.07 % and thus increased to 81.88%. Subsequently, according to Landis and Koch [1977] (Sec. 4.4.1, Fig. 4.5), κ is now ranked as almost perfect agreement and the result is in general acceptable.

LC Class	**Buildings**	**Roads**	**Trees/ Bushes**	**Meadow**	**Water**	**Bare Soil**	**Sum**
Buildings	287	20	3	0	0	21	331
Roads	38	68	0	1	0	22	129
Trees/Bushes	3	3	389	3	0	4	402
Meadow	6	0	4	65	0	6	81
Water	0	0	3	0	35	0	38
Bare soil	0	3	0	0	0	72	75
Sum	334	94	399	69	35	125	1056

Table 5.3.: Confusion matrix of classification postprocessing: Reevaluation with 1056 reference instances. Diagonal highlighted in color: Correctly classified instances.

In particular the classes “roads” and “buildings” benefit from the application of post-classification rules (Fig. 5.3). The “buildings-roads” and “roads-buildings” confusion is improved, whereby more buildings are assigned to the class “buildings” and less build-

ings are assigned mistakenly to the class "roads", and the other way around. The "bare soil-buildings" confusion can be enhanced by classifying less bare soil objects on the map mistakenly as "buildings". Yet, the improvement is made at the expense of the class "roads". These shifts are clearly illustrated when examining the PA and UA of postprocessing (Fig. 5.4).
PA and UA raise for "buildings" by 1.2% to 85.9% and by 2.8% to 86.7%, respectively. "Roads" show an improvement of 3.2% in PA, raising to 72.3%, but a degradation of UA by 1.1% (Tab. 5.4). Thus, more roads on the ground have now been classified as such, yet, at the same time, the probability that a road labeled as such is actually a road on the ground, decreases slightly.
The rules were applied based on the assumption that if an object of one class is entirely

LC class	PA	UA
Buildings	85.9%	86.7%
Roads	72.3%	52.7%
Trees/Bushes	97.4%	96.7%
Meadow	94.2%	80.2%
Water	100%	92.1%
Bare soil	57.6%	96%

Table 5.4.: Producer's and User's accuracies of classification postprocessing

nested inside objects of another class, it most probably belongs to the class surrounding it (Sec. 4.2.5). The relatively high commission error of roads in the classification task indicated a fair amount of included foreign objects in the class "roads" beforehand and therefore explains the confusion in the post-classification task. Thus, it can also happen that a correctly classified object surrounded by misclassified objects of another class will be assigned to that class mistakenly. This highlights the challenge of post-classification methods, performing in one area well and in others not (trade-off). "Bare soil'" could be improved slightly (Tab. 5.4), but in summary, the most effective improvement could be achieved for the class "buildings".

Beyond the statistical classification accuracy, a visual assessment of the results is accomplished (Fig. 5.2). The resulting LULC map shows realistic class assignments and clearly reveals the structure of the city. Noticeable misclassifications occur in the Southern fringe of the city, where a parking area and sealed surfaces of industrial use are classified as buildings. Buildings are also frequently confused with rocks or cliffs along the coast, especially in Western parts of the city. On a smaller scale, the confusion within the classes "roads", "buildings" and also "bare soil" is clearly recognizable. Most notably, they involve small IOs that are assigned to the wrong class due to very similar characterist-

ics. Basically, difficulties arise from the segmentation phase, where IOs are generated by minimizing their average heterogeneity (Sec. 4.2.1). In the case of very similar areas that nevertheless belong to different classes, IOs happen to be generated comprising parts of different classes. These "mixed" IOs constitute problems in the classification phase, since they imply mixed information about two classes and are easily assigned to the wrong class. In particular, this phenomenon is observed in the urban environment concerning the classes buildings and roads (Fig. 5.2).

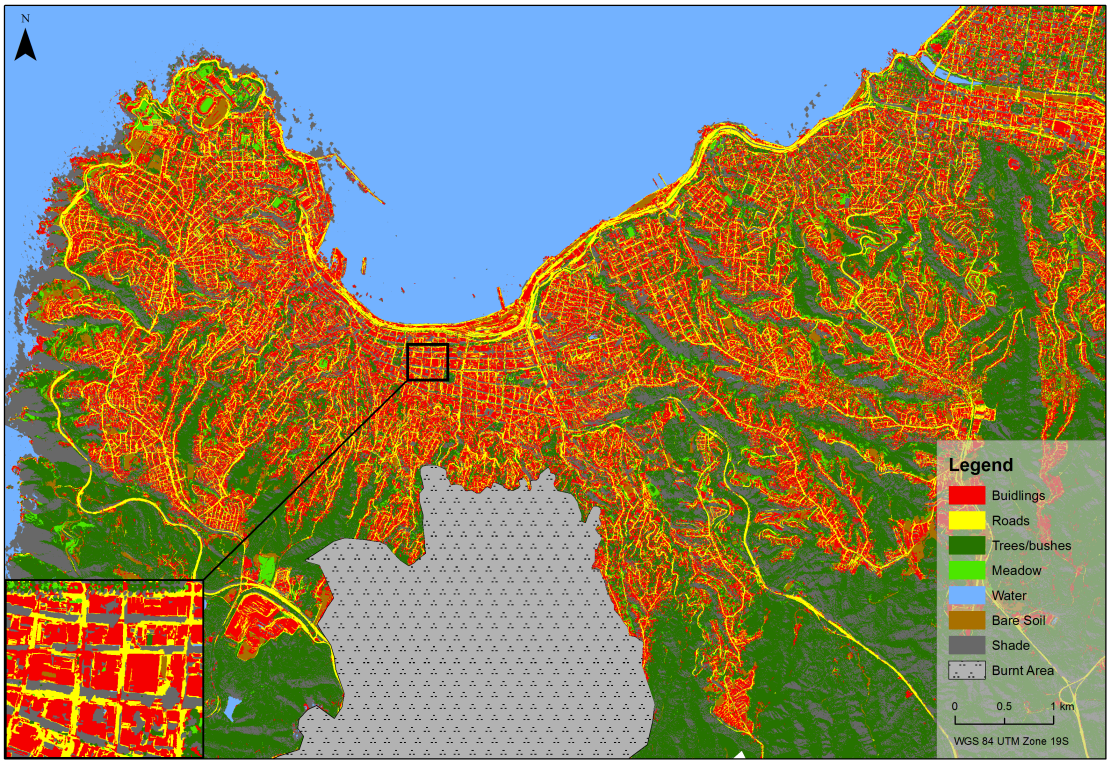


Figure 5.2.: LULC Map: Classification was carried out by means of the rotation forest classifier ensemble. OSM-data served as basis for automated generation of training samples. Burnt areas were excluded from the analysis.

5.2. Exposure Estimation

5.2.1. Number of Buildings

Preparing Dataset for Estimation of Number of Buildings

By means of a new segmentation method, it was aimed at generating image objects (IOs) of the size of buildings in order to detect single separate buildings. In addition to the multispectral channels, information from edge detection and morphological filtering

was considered in the segmentation process. However, the segmentation for building detection remains a difficult task since buildings reveal very diverse appearances in terms of their size and shape as well as their material and therefore reflectance in VHR images. As figure 5.3 shows, Valparaíso consists of buildings of very different sizes, from small wooden huts in hilly regions to huge commercial or apartment buildings in the central and harbor area. Therefore, a single scale parameter for the whole study site is not capable of capturing the range of all building sizes. Thus, bigger buildings often consist of more segments, especially if they have varying roof structures. Varying materials, structures and parts of rooftops (e.g., chimneys, windows, etc.) complicate the segmentation and its interpretation, since every variety in suchlike structures leads to differences in reflection and thus are covered by single image objects, respectively.

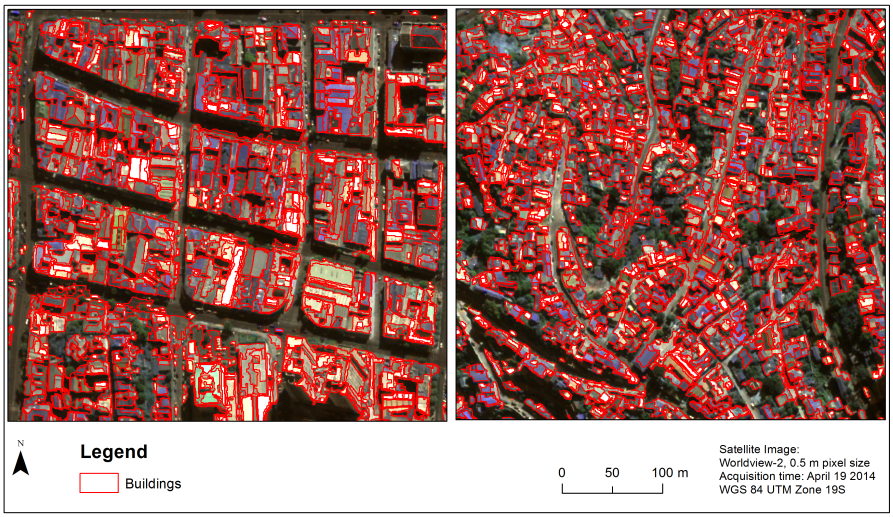


Figure 5.3.: Image Segmentation for well-structured harbour area (left) and more complex hilly area (right). Multi-resolution segmentation was conducted considering jointly 4 multispectral bands, additional bands obtained with morphological operators (erosion, dilation, opening, closing with a SE-size of 5 pixels) and canny and lee sigma edge detection for the pan channel

These varieties also have a strong influence on the applied edge detection algorithms that are sensitive to sudden changes in local intensities. Edge detection filters have shown promising results in well-structured built-up areas with detached houses. However, Valparaíso demonstrates complex and dense urban structures with blocks of buildings being attached to each other. In these areas, it is difficult to detect intensity changes on edges and the differentiation of single buildings becomes a major problem. In particular, this is due to the fact, that the algorithms detect too many edges as a consequence of high variable structures within single buildings (Fig. 5.3). However, the segmentation approach performs satisfactorily in terms of building detection despite the discussed difficulties and subsequently forms a reasonable basis for building exposure estimation.

Multivariate Regression Analysis

In terms of estimation of the number of buildings within the study site, it was aimed at reproducing the relationship between a set of extracted features and the number of buildings by three different regression methods. Therefore, descriptive features were computed for prior delineated meaningful IOs, making use of landscape structure metrics (LSM) (Sec. 4.3.1, Fig. 4.4).
In terms of the descriptive LSM, several difficulties influencing the performance of all methods can be observed:

- Measures strongly depend on the underlying spatial unit
- The number of IOs per real-world building depends on its substructure, rather than its size
- Several metrics are ambiguous on borders of cities, where only few IOs fall within a grid cell

When describing landscape patterns, the choice of the spatial unit is crucial. A trade-off has to be made in order to be able to capture certain structures (unit should not be too small) and differentiate between diverse structures at the same time (units should not be to too big). Thereby, throughout several experiments, the unit of 250×250 m^2 sized grids was found appropriate to reflect the object structure.
Concerning the regression analysis, table 5.5 depicts the model's fitting error values for the three conducted methods. While MAE and RMSE analyze the average difference in observed and estimated values (absolute fit of the model to the data), the relative error measures (RAE, RRSE and MAPE) divide those differences by their variation (Sec. 4.4.2). Therefore, they serve well for comparing the performance of the different models. The correlation coefficient (R) provides acceptable results in all three methods, suggesting a substantial correlation between the independent and dependent variables (Sec. 4.4.2). However, R by itself does not guarantee that the model fits the data well. Therefore, further error measures are analyzed in order to account for the influence of outliers and the number of variables (Tab. 5.5).
The mean absolute deviation of estimated buildings from the validation dataset ranges from 35 houses (GPR), 26 houses (LR) to 21 houses (SVR) for each grid cell. This corresponds a mean deviation of 60.3%, 35.7% and 26.2% relative to the observed values, respectively (MAPE). Furthermore, the RMSE is an important criterion to assess the quality of fit of an prediction model. SVR provides the lowest value here with 27.97, indicating the best fit, while GPR shows a quite high value. LR reveals the greatest variance in the individual errors in the application (difference MAE and RMSE). Yet, rather small differences between MAE and RMSE in general reveal rather insignificant

Error	LR	GPR	SVR
R	0.82	0.87	0.90
MAE	26.44	35.76	21.61
RMSE	37.08	42.64	27.97
RAE	45.96%	62.16%	39.29%
RRSE	55.80%	64.18%	43.53%
MAPE	35.72%	60.31%	26.21%
Prediction	72,412	85,111	64,803

Table 5.5.: Model's fitting error values: Correlation coefficient (R), mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE), root relative squared error (RRSE), mean absolute percentage error (MAPE) and overall prediction of number of buildings in Valparaíso. Tested regression approaches: Linear Regression (LR), Gaussian Processes for Regression (GPR) and Support Vector Regression (SVR).

influence of outliers, respectively, large deviations. In context of relative error measures, the MAPE gives a good overview on the results. SVR features the smallest deviation between actual and predicted values of 26 %. With a MAPE of more than 60%, GPR proves to be unsuitable for this study.

Given the validation results, SVR provides best accuracies with biggest correlation and smallest error estimates. GPR provides a quite high R but poor relative error measures. Concerning the respective estimated number of buildings, the output of SVR shows a difference about 20,000 buildings to GPR, and 7,000 buildings to LR. Taking into account the error measures, it can be observed that GPR heavily overestimates the number of buildings.

While the non-parametric methods (SVR and GPR) are conducted concerning all 12 attributes, for LR an attribute selection is applied in order to eliminate collinear attributes (see Sec. 4.3.1). Thereby, it is now interesting to examine which features are chosen to be appropriate and meaningful by the algorithm. The generated regression model reveals not only chosen features but also their influence on the function for the determination of the number of buildings (NB) (Eq. 5.1).

$$\begin{aligned} \text{NB} = &(-6.6356 * \text{Mean Segment size}) + (11.8992 * \text{Standard Deviation}) \\ &+(-8.5522 * \text{Variation}) + (149280.52 * \text{Relative Number of Segments}) \\ &+(-0.0097 * \text{Edge density}) + (-240.2959 * \text{Shape Index}) + 1053.3034 \end{aligned} \tag{5.1}$$

Negative coefficients signify a negative effect on the number of buildings and a positive coefficient a positive effect. The strongest influence on the output has the relative number of segments per grid cell, resulting in a higher number of buildings, which is quite reasonable. The other measures refer to complexity, fragmentation and variation of segment size and shape. They are not immediately graspable, since segments cannot be simply

linked to real world objects (buildings) due to the high resolution of the satellite imagery. However, in the resulting regression model a small mean segment size has a positive effect on the number of buildings. The smaller the segments, the more segments form one building on the ground. Furthermore, a decreasing shape index (indicating more complex rectangular segments) suggests more buildings per grid.

The analysis of the LR output already gives an idea of the capabilities of the descriptive variables, since its performance is promising for a basic regression method, taking into account only linear relations (Tab. 5.5). However, it becomes obvious that performances of supervised regression approaches depend strongly on the quality and quantity of the data used to train the regressor [Pasolli and Melgani 2011]. In this work, the number of training samples is limited (see section 4.3.1).

LR and SVR partly output unrealistic negative or super-positive prediction values within grid cells of poor segment coverage at borders of urban agglomerations. Okujeni et al. [2013] deal with similar problems when using SVR. They assume that this phenomenon indicates estimations beyond the known training space and therefore leads to improper extrapolation [Okujeni et al. 2013]. Thereby, it has to be considered that available OSM-data serving as training samples are not only limited concerning quality, quantity and completeness, but are moreover spatially highly aggregated (see Sec. 3.2.2). Hence, the training data is not unconditionally representative. The data represents predominantly the city structures within the HOT OSM areas, whereas LSM in other areas of the city are possibly diverging. In addition, using SVR, ambiguity of training data concerning the features per cell can lead to interpolation uncertainties when fitting the hyperplane. These errors lead to reduced mapping performances, namely inaccurate building number estimates or unrealistic estimates, e.g. negative or super-positive values [Okujeni et al. 2013]. By excluding these outliers from the results of both methods, meaningful and realistic results are achieved. Moreover, a possible error source for GPR and SVR is overfitting of the model. Overfitting means an overmatching of the resulting function on the training data [Campbell and Wynne 2011]. This results in a decreasing error in terms of the training data, but bad performance is obtained when making test predictions [Rasmussen and Williams 2006].

Since OSM data is limited, there is a limited number of observations available relative to the available parameters describing the building patterns. This can result in poor predictive performance of the model, as it exaggerates minor fluctuations of the data due to the description of random errors or noise rather than the underlying relationships in the model. Since subsequent attribute selection for GPR and SVR resulted in more erroneous predictions, it is possible that few parameters in this case are not meaningful enough or show poor correlation with the number of buildings solely. Bazi and Melgani [2010] emphasize the potentially poor performances and the loss of computational effi-

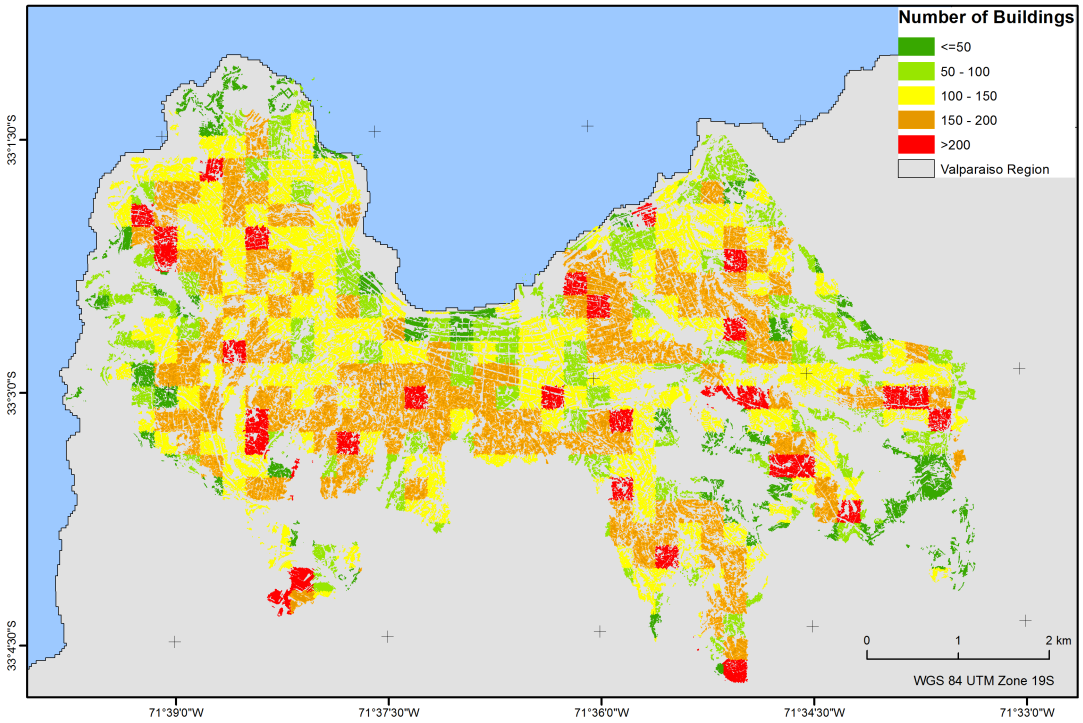


Figure 5.4.: Number of buildings: Estimation retrieved by Support Vector Regression (SVR) for 250×250 m^2 grid cells, based on OSM-data and Landscape Structure Metrics (LSM)

ciency of GPR in high dimensional spaces. Dimensionality in this cases does not seem high, but as discussed above, in terms of the number of available training samples it can cause difficulties.

Figure 5.4 illustrates the building estimations resulting for SVR. The SVR output shows promising results in terms of both, statistical and visual assessment. In the harbor area and city center few buildings were estimated which corresponds to visual interpretations. Moreover, realistic low estimations at the city borders are displayed, except for some outliers in the South of the city. Additionally, high estimated building numbers that occur in particular in the hill region seem realistic. For the study site, SVR was capable to deal best with the complex dataset, revealing realistic results. Once more, the performance quality for complex nonlinear problems and robustness of support vector machines [Camps-Valls et al. 2006, Chen et al. 2010, Tuia et al. 2011] is confirmed.

5.2.2. Estimated Period of Construction

In order to evaluate the urban expansion of Valparaíso, LANDSAT classifications from 1975 to 2010, processed at the German Aerospace Center (DLR), are used. The built-up

masks (see Sec. 3.1, Fig. 3.2) were retrieved from LANDSAT MSS 1975 with a resolution of 60 m, from LANDSAT TM 1990 and 2000 with 30 m resolution and LANDSAT ETM+ 2010 with also 30 m spatial resolution. Since the landuse classification retrieved from Worldview-2 in this study has a spatial resolution of 0.5 m, it is resampled to a 30 m resolution in terms of comparability.
Figure 5.5 and table 5.6 illustrate the increase of the built-up area from 1975 to 2014, as well as the development of the building stock that was estimated with the help of SVR (Sec. 4.3.1). The outcome shows a monotonous increase of the number of buildings and

Date	Built-up area (km^2)	Number of buildings
1975	18.07	42,807
1990	20.5	48,645
2000	23.6	55,925
2010	27.4	65,102
2014	27.2	64,610
Loss forest fire	1.2	2,815

Table 5.6.: Estimated period of construction 1975-2014: Approximation of the development of built-up area and affiliated number of buildings estimated based on OSM and SVR

the built-up area in the course of time. However, an irregularity can be observed from 2010 to 2014, where the number of buildings shows a decrease instead of an increase. This effect is connected with the 2014 forest fires that affected various districts of the city (Sec. 3.1). The time series of Landsat imagery of Valparaíso therefore permits an analysis of the built-up areas and affiliated buildings affected by the forest fire in April 2014. A computational comparison of classification areas of 2010 and 2014 (before and after the forest fires) reveals that 2,815 buildings were affected during the 2014 fires compared to 2010. This amount of buildings matches the amount officially published in Chile [ONEMI 2014]. Ignoring the affected buildings, the analysis reveals that from 2010 to 2014 the number of buildings still increased by 2,322 until 2,815 buildings were burnt during the disastrous event (Fig. 5.5).

5.2.3. Estimation of Population Statistics

The selection of meaningful error measures in order to evaluate the population disaggregation methods conducted in section 4.3.2 is a challenging task that has to be tackled carefully analyzing precisely the given data. The complex urban structure and usage differences in Valparaiso with respect to residential and non-residential usage types require the assessment of the results in a spatial context and geographically differentiated.
Several error measures have been tested and analyzed in order to best assess the performances of the methods (see Sec. 4.4.2). Relative measures are suggested to be most

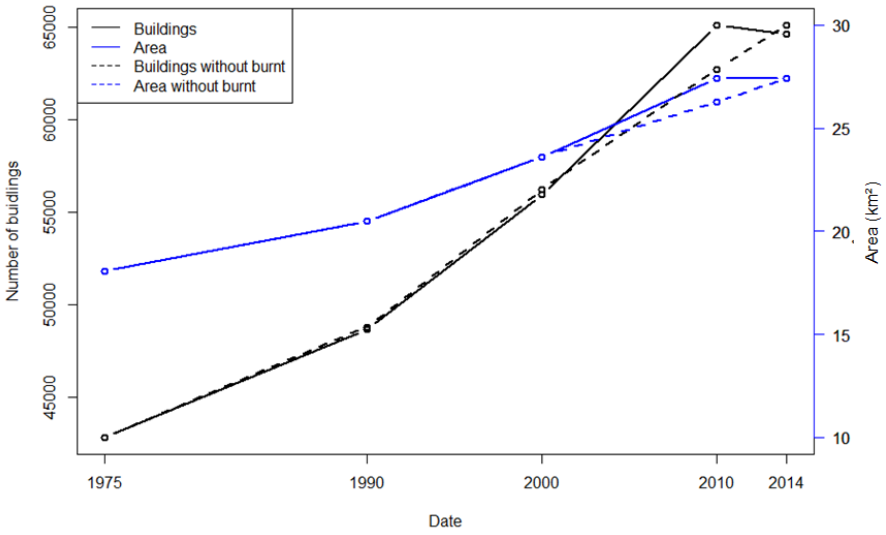


Figure 5.5.: Estimated period of construction 1975-2014: Built-up area and estimated number of buildings including in April 2014 burnt areas compared to excluding burnt areas.

appropriate to ensure comparability within and between different methods. However, the underlying data has to be suitable to assess results in this way. The observed population sizes per building blocks range between 0 and 728 with a mean population size of 74. Thus the underlying reference data is of low-volume (small numbers). Since relative measures like the MAPE or its modification SMAPE (see Sec. 4.4.2) tend to show an error relative to its reference, the denominator can be quite small or even zero in building blocks of low population size or complete non-residential use. This leads to a steep increase of the measure right up to undefined errors. Yet, building blocks without population (size = 0) cannot be excluded from the error calculations since they still hold important information and especially depict areas where ancillary information about the usage type can contribute valuable improvement. Armstrong and Collopy [1992] furthermore argue that the MAPE "puts a heavier penalty on forecasts that exceed the actual than those that are less than the actual value". Thus, the MAPE puts a heavier penalty on negative errors (when $x_i < \hat{x}_i$) than on positive errors. Thereby, especially misestimations for building blocks with very low population sizes are heavily penalized, leading to big errors and easily dominate the overall statistics, which is undesirable. Thus, when working with low-volume data the MAE is proposed as convenient measure of error in the literature, while the MAPE and other percentage-based statistics are suggested to be avoided [Chai and Draxler 2014, Makridakis and Hibon 1995]. Yet, the MAE as an absolute measure is less meaningful for global interpretations as relative measures. In this study, this disadvantage is addressed by outlining and interpreting the error results in a differentiated and spatial context (Fig. 5.6, 5.7). Therefore, the root mean square error (RMSE) and the

Disaggregation Method	RMSE	MAE
without ancillary information		
Simple Areal Weighting (SAW)	104.67	46.14
with ancillary information landscape metrics		
Number of buildings (NB)	81.74	42.69
Dasymetric mapping built-up density (BUD)	89.77	51.61
Dispersion Index (DI)	114.94	65.26
Area-Weighted Mean (AM)	96.42	52.16
with ancillary information OSM		
SAW OSM 1 class SL	104.49	44.73
SAW OSM 2 classes SL	103.6	43.37
SAW OSM 1 class BL	79.72	41.59
SAW OSM 2 classes BL	80.06	40.31
BUD OSM 1 classes SL	82.56	48.63
BUD OSM 2 classes SL	83.36	49.14
BUD OSM 1 classes BL	79.25	48.01
BUD OSM 2 classes BL	79.31	48.14
NB OSM 1 class SL	80.15	41.80
NB OSM 2 classes SL	80.12	41.77
NB OSM 1 class BL	86.93	45.85
NB OSM 2 classes BL	79.71	42.94

Table 5.7.: Overview population disaggregation: RMSE and MAE are displayed for all conducted methods. SL: Segment Level, BL: Block level

mean absolute error (MAE), which are also well-known accuracy measures in the field of population density estimation [Bakillah et al. 2014, Liu et al. 2008, Langford 2004], were captured (see Tab. 5.7). Thereby, it is aimed at assessing the results not only concerning their overall performance (Tab. 5.7), but also concerning the error origin (Fig. 5.6) and their spatial context (Fig. 5.7). This is due to the fact, that the overall results still face the problem of extremely big errors respectively outliers that influence the statistics.

Table 5.7 gives an overview on the performances of the deployed methods. The difference between RMSE and MAE varies between the different methods, indicating different variances in the individual errors (the greater the difference, the greater the variance). The deviation of estimated and reference values is more balanced when the RMSE is smaller, i.e., there are less great errors for single buildings blocks. This is emphasized in Fig. 5.7 where especially the marginal areas are higher overestimated (pink areas) by SAW than e.g. by BUD method.

The performance of methods utilizing LSM have several difficulties concerning the underlying unit and complex structures which has been discussed in section 5.2.1. The building blocks are not only quite small but also vary in their size (small BB towards the center, big BB at city edges) and therefore landscape metrics experience difficulties to display representative and meaningful structures. Furthermore, especially city borders where BB are not completely filled with image objects, lead to confusion and misestimations. In the

case of density calculation for example, these areas are misleading, indicating low density regions where there might be high dense structures. The same applies for the DI and the AM, that are not appropriate to describe settlement dispersion for such small units.
Thus, the enhancements of the density-based population disaggregation method by the Dispersion Index (DI) and the Area-Weighted Mean (AM) cannot outperform the density-based approach. Especially the DI shows a poor performance with the highest RMSE and MAE. The model was developed by Wiesner and Taubenböck [2014] for the analysis of the settlement pattern of a whole city. Although an adjustment was conducted, it could not be fitted well for the small spatial units of the building blocks for Valparaiso.

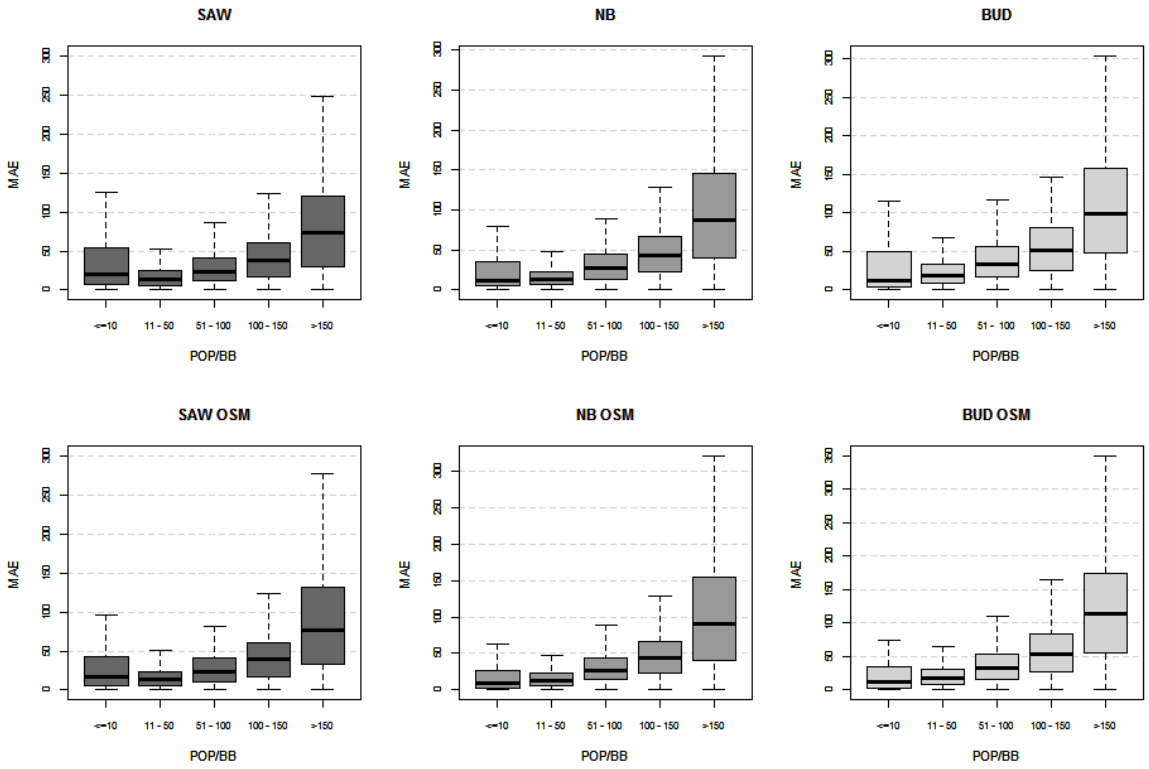


Figure 5.6.: MAE for population disaggregation based on Simple Areal Weighting (SAW), Number of buildings (NB) and built-up density (BUD), and their OSM enhancements, respectively. The box represents the interquartile range (50% of all values between the first and third quartile), the black line inside the box represents the median, the whiskers display the 1.5-fold interquartile distance.

For a more detailed analysis concerning the error relations with respect to the size of the building blocks, the MAE is chosen as suitable error measure (see Fig. 5.6). The RMSE was not found to be appropriate, since it varies with the variability within the distribution of error magnitudes and with the square root of the number of errors as well as with the MAE [Res et al. 2005]. The MAE is a more natural measure of average error and unlike the RMSE it is unambiguous and therefore easier to realize and interpret. Boxplots are

a suitable method to chart and interpret the distribution of the errors, especially since outliers are troublesome when comparing the quality of different methods based on the overall accuracy [Armstrong and Collopy 1992]. In addition to the SAW, the two best performing methods using ancillary data are taken to illustrate the MAE per BB population sizes and the effect of integrating OSM information.

Figure 5.6 features a similar pattern indicating similar error structures throughout the plotted methods and their OSM variations. Low populated and especially high populated areas seem to be prone to misestimations. The range of dispersion of MAE in high populated areas is extremely large, not only for the intermediate 50% of data (box) but also for the interquartile distances (whiskers). Errors in high populated areas do not seem to follow a specific pattern and can be quite low and extremely high depending on different reasons for misestimations, e.g. the degree of multistorey buildings in a BB. The methods based on BUD, NB and AM can outperform the SAW concerning RMSE indicating a decrease of deviations in difficult areas (Tab. 5.7). Concerning the MAE, the method based on NB performs better than the SAW in lower populated areas but poorer in high populated areas which ascribes to difficulties experienced beforehand in section 5.2.1. The method based on BUD slightly outperforms the SAW in low populated areas, but amplifies errors in high-populated areas, which can be ascribed to low-dense multistorey building colonies. The inclusion of OSM data shows slight improvements within all methods concerning the RMSE and the MAE, respectively (Tab. 5.7). The MAE for low populated areas can be reduced for all three methods by integrating OSM data. Overestimation can be mitigated to a certain extent since the integrated OSM data reveals some areas of non-residential use (Fig. 5.6). OSM information are concentrated in the harbor area or city center, therefore alterations are especially expected in this region. Yet, underestimations in high populated areas cannot be prevented. They most probably arise from multistorey apartment buildings and skyscrapers. Based on the available data, this problem could not be addressed so far and further information is needed (e.g. normalized Digital Surface Model (nDSM)). Eventually, it should be noted, that aerial interpolation is always also subject to errors from original aggregation (the population number that is disaggregated). The performance of all interpolation methods depends largely of how source zones and target zones are defined, the degree of generalization in the interpolation process, and the characteristics of the partitioned surface [Wu et al. 2005]

Fig. 5.7 illustrates the population distribution conducted by means of the three best performing methods with OSM extension and gives an impression of over- and underestimations in a spatial context. A similar pattern of overestimation can be observed throughout the outcomes of all three approaches (Fig. 5.7, a, b, c). Areas of overestimations are mainly located at the waterfront near the center, the cemetery and large sports

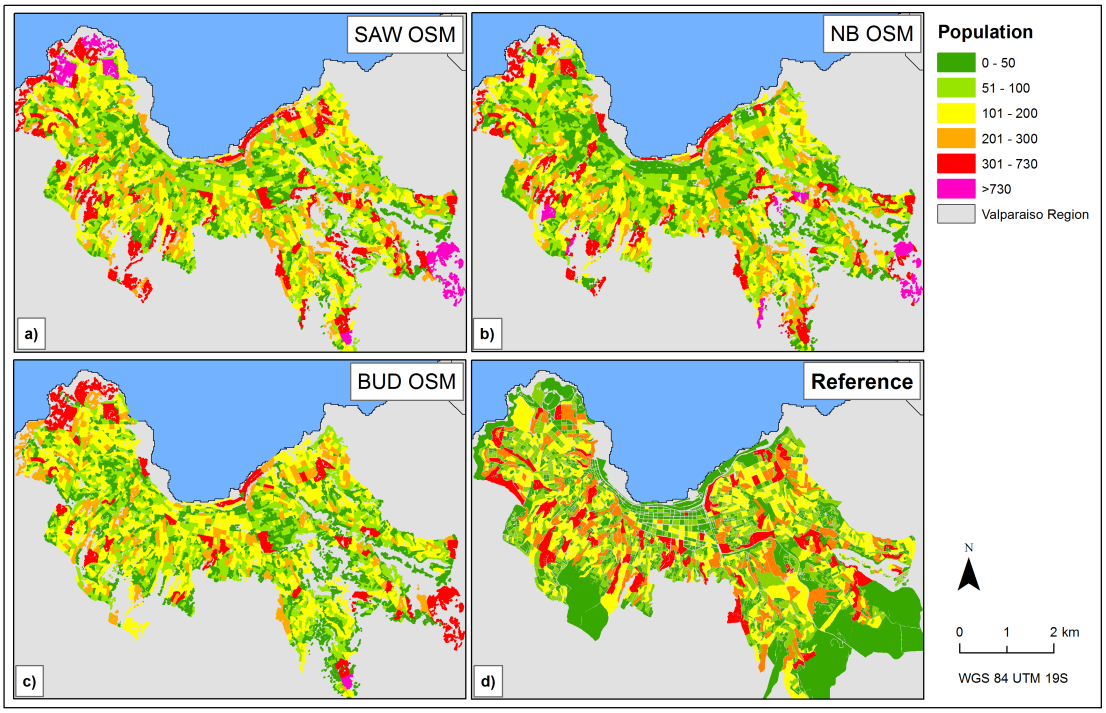


Figure 5.7.: Population Disaggregation enhanced with OpenStreetMap data based on a) Simple Areal-Weigthting (SAW), b) Estimated number of Buildings (NB) and c) built-up density (BUD). d) Census 2002 Reference. Pink areas indicate an estimation exceeding the maximum number population per building block in the reference data

and health area in the Northwest, the airfield in the Southeast, an industrial plant in the South of the city as well as various misclassified built-up areas like parking lots or bare soils at city borders that were mistakenly classified as buildings. SAW reveals the largest divergences here, revealing particularly high population numbers in non-residential areas (Fig. 5.7, a).

The spatial allocation based on built-up density overestimates the harbour and city area more than other methods due to high built-up density but few residents. In contrast, the method based on the prior estimated number of buildings shows the best results in this area. Areas of underestimation seem to appear adjacent to uninhabited valleys between the built-up hills pointing out the difficulty of marginal areas. Furthermore they can be attributed most probably to buildings with several storeys as discussed earlier.

5.3. Tsunami Scenario

Finally, the practical utility and potential of in this study generated exposure information is illustrated in the scope of a tsunami scenario for Valparaíso. In this context, the Chilean Navy Hydrographic and Oceanographic Service (SHOA) elaborated a map defining areas that potentially could be inundated by an extreme tsunami. It shows an "extreme scenario" for an earthquake of major magnitude and is primarily based on the earthquake of 1730. The earthquake had an estimated magnitude of 8.7 and triggered a major tsunami that inundated the lower parts of Valparaíso [SHOA 2015].

Affected Asset	Amount
Buildings	2927
POIs	182
Population	9248

Table 5.8.: Estimated affected assets for a tsunami scenario

The information for the scenario development was gained through a numeric simulation (COMCOT) based on topographic, batymetric and seismic data. Thereby, the seismic parameters were estimated for the 1730 event. In Fig. 5.8, two historical places (Iglesia de la Matriz and iglesia de la Merced) that were affected by the tsunami in Valparaíso in 1730 are illustrated. By means of suchlike information the extent of the historical tsunami could be reconstructed.

By a simple overlay with in this study generated information (see Sec. 5.2), threatened and potentially affected assets can easily and quickly be derived for the exposed area. In Valparaíso, in particular the harbor area and flat parts of the city center would be affected. Since the majority of the city is located on hills (see Sec. 3.1), it is less endangered by tsunamis. In contrast, in the adjacent city of Viña del Mar, major parts of the city would be inundated due to its location in rather flat terrain (Fig. 5.8). Moreover, for Valparaíso, number of buildings, population and OSM-POIs can be extracted for the specifically exposed area (Fig. 5.8). The exposure data was generated at fine scale and is thus capable to provide information for any given area of interest. The performed categorization of available POIs furthermore permits to draw a conclusion with respect to exposed functions: More than half of the POIs account for the categories gastronomy, traffic and shopping (also see Sec. 4.3.2, Fig. 4.14).

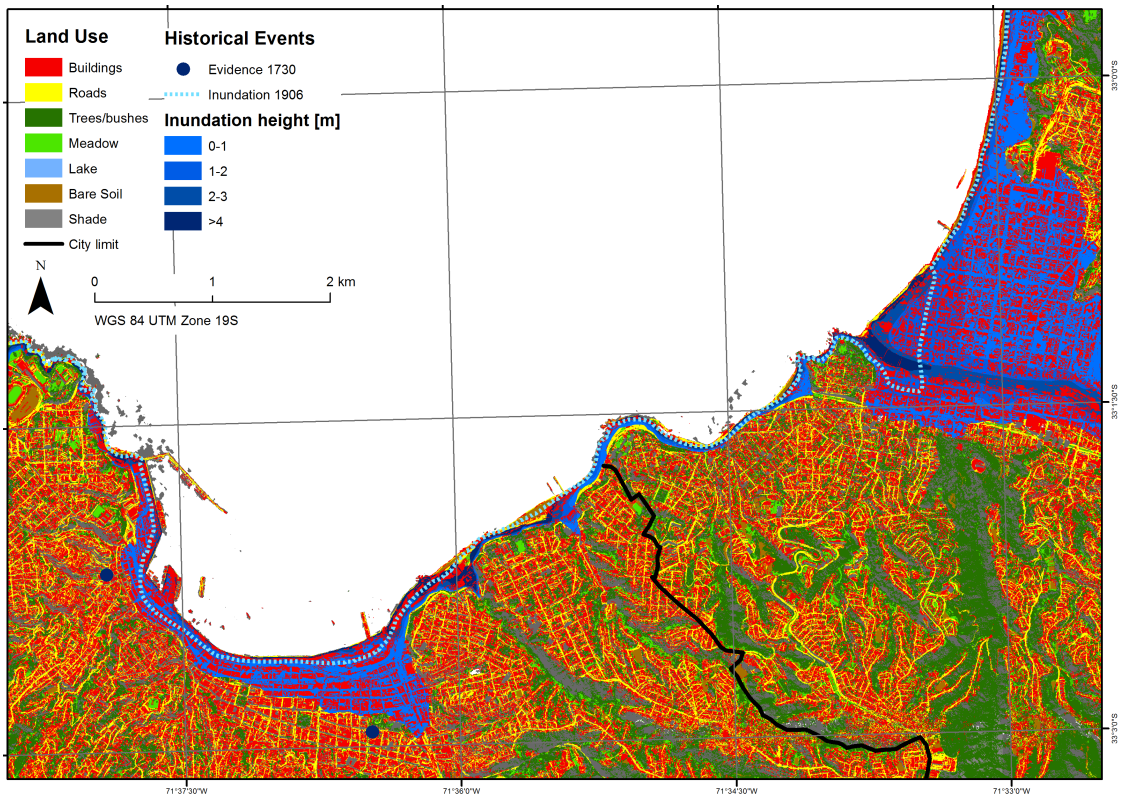


Figure 5.8.: Tsunami Scenario for Valparaíso: Tsunami scenario modelling was conducted by SHOA [2015]. Results for potential inundated areas are overlaid with in this study obtained classification results (Sec. 5.1.2)

5.4. Discussion

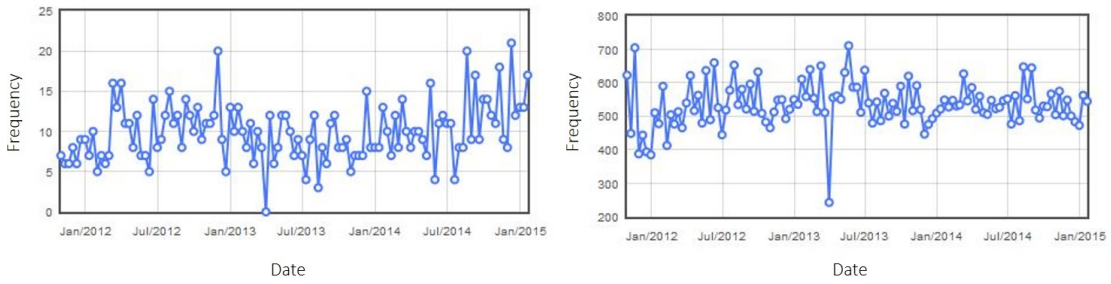
In this study, a methodology was developed in order to combine VGI data with very high resolution satellite data for the derivation of exposure information. Thereby, based on an intrinsic quality assessment, as well as the introduction of several plausibility rules, the reliability of the OSM data serving as training samples was ensured. In this manner, the RTF classifier model could be trained achieving promising classification results. In terms of exposure estimation, the number of buildings could be estimated with acceptable accuracy. However, the assessment of human exposure revealed some difficulties, suggesting that further ancillary information is necessary.

VGI Quality Issues In the area of VGI, the issue of spatial data quality is a clear challenge and it is crucial to consider quality aspects when utilizing it [Haklay et al. 2010]. In this sense, it has to be differentiated between internal and external quality [Sester et al. 2014]. Internal quality consists of the five aspects attribute accuracy, positional accuracy, temporal accuracy, logical consistency and completeness. In contrast, external

quality considers whether a dataset is suitable enough for a specific purpose, the so-called "Fitness for Use" [Sester et al. 2014]. Depending on the purpose of the data usage, quality measures have to be introduced. In the field of Open Street Map several studies towards internal data quality have been conducted. The majority evaluates the accuracy of OSM against reference sources, namely national mapping agencies [Ather 2009, Fan et al. 2014, Haklay 2010, Hecht et al. 2013, Helbich et al. 2012, Kounadi 2009, Mooney et al. 2010, Zielstra and Zipf 2010]. Other studies develop advanced quality indicators based on the OSM-history, using the behavior of the community and edits over the course of time as indicator for data quality [Neis and Zipf 2012, Barron and Neis 2013, Ciepłuch et al. 2011, Arsanjani et al. 2015]. Moreover, some studies even suggest further factors like e.g. income as influence of data quality [Neis et al. 2013]. Suchlike approaches are in particular important if no reference data is available to compare with the given VGI data. This is also the case in this study. Here, VGI is utilized with the aim of being independent of further ground-truth data.

In the Open Source Community, "Linus' law" suggests that the more users are involved in the mapping process, the more increases the control mechanism and subsequently the data quality [Mooney and Corcoran 2012, Haklay et al. 2010, Hardy et al. 2012, Arsanjani et al. 2015, Ciepłuch et al. 2011, Neis and Zipf 2012]. This suggest that manually quality checking is usually based on other users in the sense of the "wisdom of the crowd" looking at maps and "hunting for errors" [Ciepłuch et al. 2011]. Haklay et al. [2010] for example have shown that the positional accuracy of a feature improves as the number of editors increases, but with an exceeding of 13 edits per feature, no further improvement is evident.

Since there is no reference data available in this study in order to assess the OSM-data, a suchlike intrinsic quality assessment, i.e. the Contribution Index (CI) according to Arsanjani et al. [2015] is computed (Sec. 4.2.3). The CI result gives an idea of the spatial distribution of quality inside the analyzed area. Yet, it has to be treated with caution: The CI only predicts the quality of features depending on the quality of the other features inside the examined area. Results can only be considered in the frame of the study area and cannot universally be regarded as qualitative features. Therefore, OSM instances labeled as good quality features are not unconditionally of high quality - they are only good compared to other features in the study area. In this context, Goodchild and Li [2012] argue that Linus' law applies best to prominent geographic facts and less to obscure facts. This concerns facts that refer to sparsely populated or unexplored areas, which only few people are interested in. The rate of contributions depends on the quantity of existing objects and number of involved users. This is a function of resident population apart from global mapping calls for humanitarian aid [Arsanjani et al. 2015]. In the scope of this study, the quality of the OSM data should therefore be viewed and



(a) No. of daily active members Chile

(b) No. of daily active members Germany

Figure 5.9.: No. of daily active OSM-members in Germany and Chile between 2012-2015. The statistics were retrieved from the free wiki world map (OSM) and processed by Altogetherlost [2014] and Neis [2015].

assessed more precisely as well as in a general manner concerning universal facts. Figure 5.9 a and b compare the number of daily active members in Chile and Germany. Compared to Germany (which is currently the country of highest OSM-activity), Chile features an extremely low number of active users. Here, population in both countries (Germany 80,620,000 inhabitants, Chile 17,620,000 inhabitants [The World Bank Group 2015]) have to be considered, which results in roughly 0.0006% of OSM-active inhabitants in both countries. Yet, relative to the number of inhabitants in both countries, Chile features only one fifth of active users in Germany (Chile: 10,572; Germany: 48,372). Consequently, low data quality in terms of Linus' law is suggested. Figure 3.4 in section 3.2.2 shows the same statistic for the study site with only one editing user in average per month. Contributions in the scope of the HOT OSM Task play an important role in the study area (Figure 3.4,3.3), which should also be considered.

Another approach to assess quality of crowd-sourced data is the "social approach" [Goodchild and Li 2012]. This refers to where the contributions come from and why they have been made (hierarchy of trusted individuals and the motivation of the contributors) [Goodchild and Li 2012]. In their analysis of OSM data for the British Isles, Mooney and Corcoran [2012] found out that for heavily edited features (at least 15 edits per feature) 84% of edits were made by only 12% of contributors.

Figure 5.10 illustrates the average number of edited features per contributor for the study site. The area of the HOT OSM Task stands out here, with more than 40 features per active contributor per cell. It becomes obvious that about 40% of all users edited less than 5 features while also 40% edited more than 40 features, whereas the latter ones concentrate in the HOT OSM area. This suggests that the normal average of edited features per contributor is increased by the incidence of the HOT OSM Task by more than 10%. Overall, 86% of the features were edited by only 19% of the 208 contributors

(40 contributors). Thereby, it also has to be considered that a lot of contributors only followed the humanitarian call and are not usual mappers in the area. The fact that edits are conducted by a small amount of active users in the study site reinforces the suggestion of poor data quality in terms of Linus' law. Looking back at the Contribution Index that

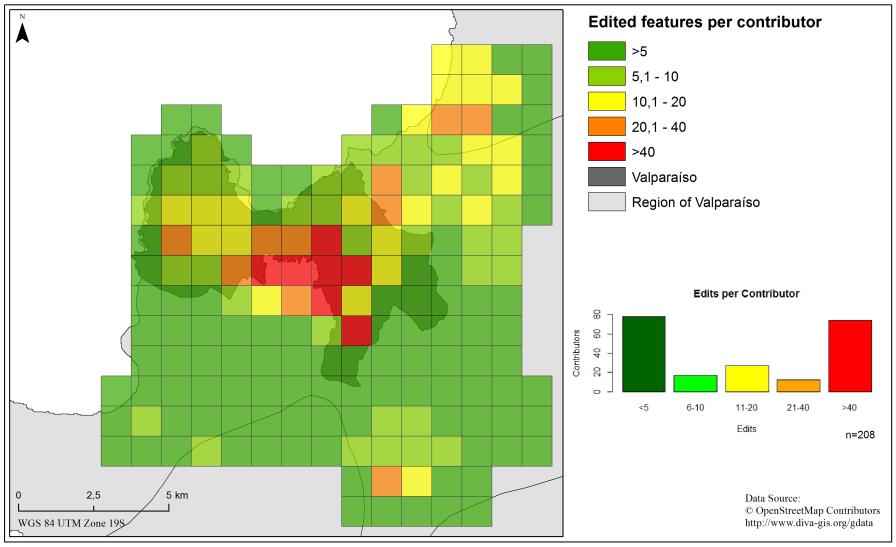


Figure 5.10.: Edits per Contributor for Valparaíso: Map: Edits that were made on average per active contributor in the the study site. Figure: Amount of Contributors that made certain number of edits.

was used in this study, it can be assumed that it does not sufficiently describe the quality of the data. OSM-based training samples require absolute positional accuracy as well as size and shape, which is not considered in the CI. As stated in section 4.1.2, the necessary positional quality for the analysis was not valid for all features and therefore a manual checking had to be conducted beforehand. However, these difficulties were additionally prevented by introducing plausibility rules in order to ensure reliability beyond the CI (Sec. 4.2.3, Fig. 4.8).

Finally, according to Sester et al. [2014], external quality measures whether a dataset is suitable enough for a specific purpose, which is termed as the "Fitness of Use". In the following the external quality and performance ability of VGI-based training samples is discussed in the scope of supervised classification and exposure estimation based on the findings within this study.

Performance Ability of VGI as Training Input for Classification When using VGI as alternative data source for remote sensing image processing, a careful selection and quality assurance process is crucial in order to ensure a high quality of the training data. As discussed earlier, several possibilities to assess VGI quality have been discussed in the literature recently, yet, current methods are still not satisfactory and need further

research.
In addition, for the selection of training samples, the completeness of available data is of interest. This is in terms of representativity rather than the necessity of numerous samples. Figure 5.1 in section 5.1.1 illustrates that the function oft the κ statistics raises monotonously with the number of training samples, but shows a plateau effect with an increasing training set size. This means, if a certain amount of training samples is reached, additional samples will not lead to further improvement. Furthermore, the distribution of available samples within the study side has an impact on the results. In this study, available OSM-data are highly spatially aggregated, which leads to an imbalance when classifying a whole scene. Illumination and image quality, as well as the material of observed assets itself can vary within an image. Therefore it is important that different characteristics are represented. Subsequently, an exhaustive, well distributed population of available data is advantageous. This is especially of concern in regions with low contribution rates, like the study site.
Moreover, it is uncertain to which extent a particular LULC class is represented by the available, underlying VGI data. In other words, the amount of OSM-features per LULC class does not necessarily correspond the occurrence of the class in reality. In the scope of this study, this issue is considered by selecting the same size of training samples for each class by stratified sampling (see Sec. 4.2.3).
In summary, the developed methodology for generating training samples comprises three working steps: A preprocessing phase, a set of plausibility rules in order to ensure reliability of the samples and stratified sampling of the final training data from areas of highest quality. The procedure facilitates the generation of training samples and achieves good classification results. Thereby, all the steps contribute to the success. Yet, it has to be considered, that the plausibility rules, the preprocessing step and quality assessment need to be applied and adjusted on a respective study site. Thereby, the preprocessing step might not be necessary for other study areas.
With respect to the choice of the classifier in terms achieved quality, the robustness of the RTF ensemble could be proven. The procedure of generating input datasets by axis rotation in order to create diverse classifiers is capable of dealing with the OSM-based input data (Sec. 4.2.4). Additionally, a traditional SVM has been trained during the experimental phase which could not achieve results at the same level, and was outperformed by the RTF.

VGI and Exposure Estimation In terms of exposure estimation, once again, the reliability of the training data is crucial and has a strong impact on the outcomes. Thereby, the completeness of OSM-buildings counted for a given spatial unit as reference is of vital importance. In this study, the reliability of the completeness of counted buildings could

not be assessed in a satisfactory manner. Furthermore, as already discussed in detail in section 4.3.1, representativity of the training input with respect to the depiction of diverse urban structures and affiliated LSM is required. The training set needs to describe and represent different existing urban structures in order to be able to estimate the number of buildings in diverse urban sceneries. Estimations beyond the known training space quickly lead to errors or inproper extrapolations (see Sec. 5.2.1). Yet, the results reveal that the generation of training samples for regression purposes by means of automatically counting OSM buildings is a promising and easily feasible approach.

Similar issues apply to the estimation of exposed population. In order to correlate usage types of areas with population densities, an exhaustive underlying VGI database is needed. This is also the case for the method conducted in section 4.3.2, where it is tried to control the population distribution by OSM-based information concerning the usage types. VGI-information about landuse as well as POIs were limited in the study site, and the developed method should be viewed as an exemplary possible approach that requires further research. Additionally, a realistic distribution of the population seems difficult based on the available data and without doubt a more complete OSM dataset as well as ancillary information is required. The issue of underestimation in areas of greater building height cannot be prevented based on the available data. This problem can so far only be addressed by integrating information on building heights. Yet, it could be shown that based on little available OSM-information, strong overestimations could be prevented in areas of non-residential use.

Since cities are composed of very different structures, it may also make sense to undertake a subdivision concerning diverse urban structure types to then conduct different disaggregation methods that fit the respective conditions best. In particular in Valparaíso, the diverse urban characteristics (e.g., flat harbor area with big buildings and hilly regions with small detached houses) influence the performances of the deployed methods that therefore revealed spatially varying qualities.

6. Conclusion & Outlook

Conclusion The main objective of this research was to investigate the potential of VGI in remote sensing data processing in order to derive exposure information. Thereby, OSM-data are deployed in three different applications: LULC classification, estimation of the number of buildings and estimation of exposed population. As far as we know, an incorporation of OSM for exposure estimation has not been performed heretofore. In general, the combination of VGI with remote sensing classification techniques has been addressed rarely until now, although its potential is gaining increasing recognition (Sec. 1.2). Even if quality standards constitute a non-negligible challenge, VGI is considered as more trustworthy than other forms of user-generated content because geographic facts are perceived to be objective and replicable [Flanagin and Metzger 2008].

In this study, a work flow is proposed, that is able to generate useful training samples for supervised landuse classification by means of the combination of OSM and remote sensing derivatives. OSM-data is not only selected by stratified sampling from predicted best quality areas, additionally, the data has to meet formulated conditions that ensure the plausibility in terms of their class affiliation. The OSM-class information is combined with the segmented image by majority vote and classified by the RTF ensemble with promising results. With an overall accuracy of 86.7% and a κ of 81.8% , the quality assessment and classification procedure shows great success and potential for further improvements.

LULC classification that is obtained in this way serves to extract additional exposure information. In terms of physical exposure, OSM-data and remotely sensed products are fused to obtain training samples for the estimation of the number of buildings. Although few available OSM data within the study site made the generation of sufficient, representative training data difficult, promising results could be achieved. Deployed machine learning regression methods turned out to be appropriate for exposure estimation, yet, the importance of representative, well distributed and highly accurate training samples must not be underestimated. Furthermore, with respect to building detection, a DSM or nDSM would be beneficial and capable of further improvement.

The problem of disaggregating population from coarse community statistics to pixel level for the provision of a fine scale population dataset in the scope of human exposure could not be solved satisfactorily in a feasible manner. Population disaggregation was con-

ducted with traditional methods and intended to be improved by incorporating OSM-information. All methods that were tested in this study reveal unsatisfactory results and it can be assumed that ancillary information is essential in order to obtain promising results. In particular a nDSM providing building heights is highlighted to be indispensable in order to prevent underestimations. In addition, supplementary landuse information is preferable, which in this work were retrieved from OSM. Despite the data scarcity in terms of landuse information and POIs, the developed approach incorporating OSM data showed that already few data are sufficient to prevent strong overestimations in non-residential areas.
In summary, the obtained results highlight the potential of incorporating VGI in remote sensing applications for information extraction and exposure estimation. The study site depicts an exemplary use case for areas characterized by data scarcity and the necessity of exploiting new data sources. The incompleteness and inaccuracy of available VGI-data was challenging, yet, its usefulness could be demonstrated. With an increasing availability of VGI, these methods will be capable of high performances. Given the importance of the temporality of ground truth information, the use of VGI could be imperative for many applications in terms of improving data collection [Kinley 2013]. However, the quality of VGI is heterogeneous and location-dependent, therefore it is recommended to check the amount and accuracy of contributions before using it for those purposes.
In contrast to common supervised methods, the incorporation of OSM data has the great advantage of transferability in terms of new global data availability as well as applicability in diverse thematic and geographical areas. Supervised approaches usually utilize site-specific training data. By utilizing worldwide available OSM-data, applications are most widely disengaged from these dependencies. This is a great added value in the scope of supervised remote sensing approaches and data provision for integrative risk models, yet, appropriate methodologies in order to assure data quality and credibility are still required.

Outlook The outcomes of this study prove the potential of integrating VGI within remote sensing processes to facilitate earth monitoring for exposure information provision. With VGI, an exhaustive data source is made available for remote sensing by the proposed methodology. Yet, further research should be conducted in order to improve the potential that emerges by fusing these two powerful data sources. This also embraces the development of automation of the single work steps, which is partially given but implies further potential for future applications.
The challenge remains to ensure the training samples to be of extremely high quality due to their great impact on the feasibility and outcome of the analysis. Future research in terms of automatic quality assessment of OSM data is urgently necessary to be able

to use the data as reliable substitute for studies like this. Thereby, spatial accuracy, especially in accordance to the respective satellite imagery constitutes a major challenge. Additionally, further methods and rules for the assessment of the credibility of the VGI-based samples could lead to increasing improvements. In terms of supporting accuracy and information content, it is also worth considering additional VGI sources, e.g. Flickr photos or Wikimapia.

In order to estimate human exposure, future work should focus on the incorporation of VGI from the beginning. Bakillah et al. [2014] made a first move and examined the potential of OSM for fine-resolution population mapping. To take consideration a step further, the exposure to hazards, and hence the spatial distribution of population is time-dependent [Aubrecht et al. 2013]. Due to its up-to-dateness and widespread availability, VGI contributions show a great potential to estimate day-and nighttime population in order to assess exposure and risk more accurately. In this context, Aubrecht et al. [2011] made first efforts. In terms of modeling risk in general, the deployed approaches could be extended to provide additional information with respect to vulnerability. Thereby, it would be interesting to examine to what extent VGI can contribute to assess vulnerability (e.g. seismic vulnerability) in conjunction with remote sensing applications.

Finally, due to the fact that the amount and quality of VGI is heterogeneous and location-dependant, it would be interesting to deploy the developed methodologies in areas of higher OSM contributions rates. Thereby, the progressive development of VGI and especially OSM needs to be observed. It features a continuous augmentation of available information sources and opportunities in terms of data processing feasibility and accuracy. It should be kept in mind that the methodology in this study was developed based on an area with sparse and inaccurate VGI contributions, which emphasizes its potential for more developed regions even more. Given remote sensing as a powerful earth monitoring tool and VGI as an exhaustive information source, the combination of both has the potential to generate and update global risk models and the approach could therefore significantly contribute to disaster mitigation and response.

Bibliography

Acar, U. and Bayram, B. (2009). Building Extraction With Morphology, *RAST 2009 - Proceedings of 4th International Conference on Recent Advances Space Technologies*, pp. 33–38.

Albertz, J. (2013). *Einführung in die Fernerkundung: Grundlagen der Interpretation von Luft- und Satellitenbildern*, 5 edn, WBG, Darmstadt.

Ali, M. and Clausi, D. (2001). Using the Canny edge detector for feature extraction and enhancement of remote sensing images, *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium* **5**(C): 2298–2300.

Altogetherlost (2014). altogetherlost.com, Available from: http://altogetherlost.com/. Last Accessed: 12 December 2014.

Alvarado, G. and Moya, J. C. (2007). División Político Administrativa y Censal 2007, *Technical report*, Instituto Nacional de Estadísticas Chile (INE), Santiago de Chile.

Armstrong, J. S. and Collopy, F. (1992). Error Measures for Generalizing About Forecasting Methods : Empirical Comparisons, *International Journal of Forecasting* **8**(1): 69–80.

Arsanjani, J. J., Helbich, M. and Bakillah, M. (2013). Exploiting Volunteered Geographic Information to ease land use mapping of an urban landscape, *Proceedings of Urban Data Management Society conference*, Vol. XL-4/W1, pp. 51–55.

Arsanjani, J. J., Mooney, P., Helbich, M. and Zipf, A. (2015). An Exploration of Future Patterns of the Contributions to OpenStreetMap and Development of a Contribution Index, *Transactions in GIS*.

Ather, A. (2009). *A Quality Analysis of OpenStreetMap Data*, Dissertation, University College London.

Aubrecht, C., Özceylan, D., Steinnocher, K. and Freire, S. (2013). Multi-level geospatial modeling of human exposure patterns and vulnerability indicators, *Natural Hazards* **68**(1): 147–163.

Aubrecht, C., Ungar, J. and Freire, S. (2011). Exploring the potential of volunteered geographic information for modeling spatio-temporal characteristics of urban population A case study for Lisbon Metro using foursquare check-in data, *International Conference Virtual City and Territory 2011. Lisboa*, pp. 11–13.

Awrangjeb, M., Zhang, C. and Fraser, C. S. (2011). Improved building detection using texture information, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS)* **38**(3): 1–6.

Baatz, M. and Schäpe, A. (2000). Multiresolution Segmentation - an optimization approach for high quality multi-scale image segmentation, *in* J. Strobl, T. Blaschke and G. Griesebner (eds), *Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium.*, Wichmann, Heidelberg, pp. 12–23.

Backhaus, K., Erichson, B., Plinke, W. and Weiber, R. (2011). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*, 13 edn, Springer, Heidelberg.

Bahrenberg, G., Giese, E. and Nipper, J. (1999). *Statistische Methoden in der Geographie 1*, 4 edn, B.G. Teubner, Stuttgart.

Bakillah, M., Liang, S., Mobasheri, A., Jokar Arsanjani, J. and Zipf, A. (2014). Fine-resolution population mapping using OpenStreetMap points-of-interest, *International Journal of Geographical Information Science* **28**(9): 1940–1963.

Barron, C. and Neis, P. (2013). Mapping the Crowd - zur Rolle der Mapper bei der Qualitätsanalyse von OpenStreetMap, *AGIT 2013 - Symposium und Fachmesse Angewandte Geoinformatik*, Salzburg, Österreich, pp. 152–161.

Bazi, Y. and Melgani, F. (2010). Gaussian Process Approach to Remote Sensing Image Classification, *IEEE Transactions on Geoscience and Remote Sensing* **48**(1): 186–197.

Beck, M. W., Shepard, C. C., Birkmann, J., Rhyner, J., Welle, T., Witting, M., Wolfertz, J., Martens, J., Maurer, K., Mucke, P. and Radtke, K. (2012). *World Risk Report 2012*, Development Works-Together for People in Need.

Bellens, R., Douterloigne, K., Gautama, S. and Philips, W. (2008). Per pixel contextual information for classification of vhr images of urban areas, *International Geoscience and Remote Sensing Symposium (IGARSS)*, Vol. 4, pp. 522–525.

Benediktsson, J. A., Pesaresi, M. and Arnason, K. (2003). Classification and feature extraction for remote sensing images from urban areas based on morphological transformations, *IEEE Transactions on Geoscience and Remote Sensing* **41**(9): 1940–1949.

Berry, J. K. (2007). *Map Analysis: Understanding Spatial Patterns and Relationships*, GeoTec Media.

Bianchini, R., Feeney, G. and Rajendra, S. (2013). Report of the International Commission on the 2012 Population and Housing Census of Chile, *Technical Report November*, International Commission, Santiago de Chile, Chile.

Blaschke, T. (2003). Object-based contextual image classification built on image segmentation, *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, 2003* **00**(C): 113–119.

Bohle, H.-G. (2001). Vulnerability and Criticality: Perspectives from Social Geography, *IHDP update* **2**: 1–6.

Bruzzone, L. and Carlin, L. (2006). A Multilevel Context-Based System for Classification of Very High Spatial Resolution Images, *IEEE Transactions on Geoscience and Remote Sensing* **44**(9): 2587–2600.

Campbell, J. B. and Wynne, R. H. (2011). *Introduction to Remote Sensing*, 5 edn, The Guilford Press, New York, London.

Camps-Valls, G., Bruzzone, L., Rojo-Álvarez, J. L. and Melgani, F. (2006). Robust support vector regression for biophysical variable estimation from remotely sensed images, *IEEE Geoscience and Remote Sensing Letters* **3**(3): 339–343.

Canny, J. (1986). A Computational Approach to Edge Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Carleer, A. P. and Wolff, E. (2006). Urban land cover multi-level region-based classification of VHR data by selecting relevant features, *International Journal of Remote Sensing* **27**(6): 1035–1051.

CEUDMT (2007). Tools and Methods for Estimating Populations at Risk from Natural Disasters and Complex Humanitarian Crises, *Technical report*, The National Academies, Washington, D.C.

Chai, T. and Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature, *Geoscientific Model Development* **7**(3): 1247–1250.

Chen, G., Hay, G. J. and Zhou, Y. (2010). Estimation of forest height, biomass and volume using support vector regression and segmentation from lidar transects and quickbird imagery, *2010 18th International Conference on Geoinformatics, Geoinformatics 2010*.

Chen, Y., Su, W., Li, J. and Sun, Z. (2009). Hierarchical object oriented classification using very high resolution imagery and LIDAR data over urban areas, *Advances in Space Research* **43**(7): 1101–1110.

Chi, M., Feng, R. and Bruzzone, L. (2008). Classification of hyperspectral remote-sensing data with primal SVM for small-sized training dataset problem, *Advances in Space Research* **41**(11): 1793–1799.

Ciepłuch, B., Mooney, P. and Winstanley, A. C. (2011). Building Generic Quality Indicators for OpenStreetMap, *19th annual GIS Research UK*.

Congalton, R. G. and Green, K. (1999). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, CRC Press, Boca Raton, USA.

Crooks, A., Pfoser, D., Jenkins, A., Croitoru, A., Stefanidis, A., Smith, D., Karagiorgou, S., Efentakis, A. and Lamprianidis, G. (2015). Crowdsourcing urban form and function, *International Journal of Geographical Information Science*.

Cutter, S. L. (1996). Vulnerability to environmental hazards, *Progress in Human Geography* **20**(4): 529–539.

Dahiya, S. and Garg, P. K. (2013). Object Oriented Approach for Building Extraction from High Resolution Satellite Images, *3rd IEEE International Advance Computing Conference (IACC)*, pp. 1300–1305.

Dare, P. M. (2005). Shadow Analysis in High-Resolution Satellite Imagery of Urban Areas, *Photogrammetric Engineering & Remote Sensing* **71**(2): 169–177.

Darwish, A., Leukert, K. and Reinhardt, W. (2003). Image segmentation for the purpose of object-based classification, *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings*, Vol. 3, pp. 2039–2041.

De Lange, N. (2013). *Geoinformatik in Theorie und Praxis*, 3 edn, Springer Spektrum, Heidelberg.

De Roeck, T., Van De Voorde, T. and Canters, F. (2009). Full hierarchic versus non-hierarchic classification approaches for mapping sealed surfaces at the rural-urban fringe using high-resolution satellite data, *Sensors* **9**(1): 22–45.

Deichmann, U., Ehrlich, D., Schmal, C., Zeug, G. and Small, C. (2011). Using High Resolution Satellite Data for the Identification of Urban Natural Disaster Risk, *Technical report*, European Union and World Bank.

Del Frate, F., Pacifici, F., Schiavon, G. and Solimini, C. (2007). Use of neural networks for automatic classification from high-resolution images, *IEEE Transactions on Geoscience and Remote Sensing* **45**(4): 800–809.

Deswal, S. (2011). Modeling Oxygen-transfer by Multiple Plunging Jets using Support Vector Machines and Gaussian Process Regression Techniques, *International Science Index* **5**(1): 28–33.

Digital Globe (2015). Content Collection, Available from: http://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-5444/9113_read-17838/. Last Accessed: 02 March 2015.

DLR (2015). CATENA, Available from: http://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-5444/9113_read-17838/. Last Accessed: 02 March 2015.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A. and Vapnik, V. (1996). Support vector regression machines, *Advances in Neural Information Processing Systems* **x**: 155–161.

Duda, R. O., Hart, P. E. and Stork, D. G. (2001). *Pattern Classification*, 2 edn, John Wiley & sons, Inc., New York.

Ebden, M. (2008). Gaussian Processes for Regression: A Quick Introduction, *Technical report*.

Eguchi, R. T., Huyck, C. K., Ghosh, S. and Adams, B. J. (2008). The Application of Remote Sensing Technologies for Disaster Management, *The 14th World Conference on Earthquake Engineering, October 12-17, 2008*, Beijing, China.

Ehrlich, D. and Tenerelli, P. (2013). Optical satellite imagery for quantifying spatio-temporal dimension of physical exposure in disaster risk assessments, *Natural Hazards* **68**(3): 1271–1289.

Ehrlich, D., Kemper, T., Blaes, X. and Soille, P. (2012). Extracting building stock information from optical satellite imagery for mapping earthquake exposure and its vulnerability, *Natural Hazards* **68**(1): 79–95.

Eicher, C. L. and Brewer, C. A. (2001). Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation, *Cartography and Geographic Information Science* **28**(2): 125–138.

Epifanio, I. and Soille, P. (2007). Morphological texture features for unsupervised and supervised segmentations of natural landscapes, *IEEE Transactions on Geoscience and Remote Sensing* **45**(4): 1074–1083.

Erikson, I. and Högstedt, J. (2004). *Landslide hazard assessment and landslide precipitation relationship in valparaiso, central chile*, Master thesis, Göteborg University.

Esch, T., Thiel, M., Bock, M., Roth, a. and Dech, S. (2008). Improvement of image segmentation accuracy based on multiscale optimization procedure, *IEEE Geoscience and Remote Sensing Letters* **5**(3): 463–467.

Fan, H., Zipf, A., Fu, Q. and Neis, P. (2014). Quality assessment for building footprints data on OpenStreetMap, *International Journal of Geographical Information Science* **28**(4): 700–719.

Farr, T. G., Rosen, P. a., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D. and Alsdorf, D. E. (2007). The shuttle radar topography mission, *Reviews of Geophysics* **45**(2): 1–43.

Flanagin, A. J. and Metzger, M. J. (2008). The credibility of volunteered geographic information, *GeoJournal* **72**(3-4): 137–148.

Florczyk, A. J., Kemper, T., Kauffmann, M. and Soille, P. (2014). VGI in automatic information retrieval from big RS data, *ICGS Conference*, Vienna, Austria.

Foody, G. (2004). Thematic map comparison: evaluating the statistical significance of differences in classification accuracy, *Photogrammetric Engineering & Remote Sensing* **70**(5): 627–633.

Foody, G. M. (2002). Status of land cover classification accuracy assessment, *Remote Sensing of Environment* **80**: 185–201.

Foody, G. M. and Boyd, D. S. (2013). Using volunteered data in land cover map validation: Mapping west African forests, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **6**(3): 1305–1312.

Freire, S. and Aubrecht, C. (2011). Assessing Spatio-Temporal Population Exposure to Tsunami Hazard in the Lisbon Metropolitan Area, *Proceedings of the 8th International ISCRAM Conference*, number May, Lisbon, Portugal, pp. 1–5.

Freire, S. and Aubrecht, C. (2012). Integrating population dynamics into mapping human exposure to seismic hazard, *Natural Hazards and Earth System Science* **12**(11): 3533–3543.

Freire, S. and Gomes, N. (2013). Advancing Environmental Noise Pollution Analysis in Urban Areas by considering the Variation of Population Expoure in Space and

Time, *29th Urban Data Management Symposium, 29 - 31 May*, Vol. XL-4/W1, London, United Kingdom, pp. 155–160.

Gaikwad, S. and Pise, N. (2014). An Experimental Study on Hypothyroid using Rotation Forest, *International Journal of Data Mining & Knowledge Management Process (IJDKP)* **4**(6): 31–37.

Geiß, C. and Taubenböck, H. (2013). Remote sensing contributing to assess earthquake risk: From a literature review towards a roadmap, *Natural Hazards* **68**(1): 7–48.

Geiß, C., Jilge, M., Lakes, T. and Taubenböck, H. (2015). Seismic Vulnerability Assessment of Urban Structures with Multisensor Remote Sensing, Manuscript submitted for publication.

Geiß, C., Pelizari, P. A., Marconcini, M., Sengara, W., Edwards, M., Lakes, T. and Taubenböck, H. (2014a). Estimation of seismic building structural types using multisensor remote sensing and machine learning techniques, *ISPRS Journal of Photogrammetry and Remote Sensing* **104**: 175–188.

Geiß, C., Taubenböck, H., Tyagunov, S., Tisch, A., Post, J. and Lakes, T. (2014b). Assessment of Seismic Building Vulnerability from Space, *Earthquake Spectra* **30**(4): 1553–1583.

Geofabrik (2015). OpenStreetMap Data Extracts, Available from: http://download.geofabrik.de/. Last Accessed: 05 November 2014.

Giacco, F., Thiel, C., Pugliese, L., Scarpetta, S. and Marinaro, M. (2010). Uncertainty analysis for the classification of multispectral satellite images using SVMs and SOMs, *IEEE Transactions on Geoscience and Remote Sensing* **48**(10): 3769–3779.

Giada, S., De Groeve, T., Ehrlich, D. and Soille, P. (2003). Information extraction from very high resolution satellite imagery over Lukole refugee camp, Tanzania, *International Journal of Remote Sensing* **24**(22): 4251–4266.

Global Resources News (2014). Mapping the Fire Disaster in Valparaiso, ChileNo Title, Available from: http://globalresourcesnews.com/p-Valparaiso-disaster-mapping. Last Accessed: 22 January 2015.

Goodchild, M. F. (2007a). Citizens as sensors: the world of volunteered geography, *GeoJournal* **69**(4): 211–221.

Goodchild, M. F. (2007b). Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0, *International Journal of Spatial data Infrastructures Research* **2**: 24–32.

Goodchild, M. F. and Glennon, J. A. (2010). Crowdsourcing geographic information for disaster response: a research frontier, *International Journal of Digital Earth* **3**(3): 231–241.

Goodchild, M. F. and Li, L. (2012). Assuring the quality of volunteered geographic information, *Spatial Statistics* **1**: 110–120.

Guo, Y., Tian, X., Li, Z., Ling, F., Chen, E., Yan, M. and Li, C. (2014). Comparison of estimating forest above-ground biomass over montane area by two non-parametric methods, *International Geoscience and Remote Sensing Symposium (IGARSS)*, Vol. 2, pp. 741–744.

Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets, *Environment and Planning B: Planning and Design* **37**(4): 682–703.

Haklay, M., Basiouka, S., Antoniou, V. and Ather, A. (2010). How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus Law to Volunteered Geographic Information, *The Cartographic Journal* **47**(4): 315–322.

Hardy, D., Frew, J. and Goodchild, M. F. (2012). Volunteered geographic information production as a spatial process, *International Journal of Geographical Information Science* **26**(7): 1191–1212.

Hecht, R., Kunze, C. and Hahmann, S. (2013). Measuring Completeness of Building Footprints in OpenStreetMap over Space and Time, *ISPRS International Journal of Geo-Information* **2**(4): 1066–1091.

Helbich, M., Amelunxen, C. and Neis, P. (2012). Comparative Spatial Analysis of Positional Accuracy of OpenStreetMap and Proprietary Geodata, *GI - Forum 2012 – Symposium and Exhibit GIScience and Technology/Learning with GI*, Salzburg, Austria.

Herold, M., Gardner, M. and Roberts, D. (2003). Spectral resolution requirements for mapping urban areas, *IEEE Transactions on Geoscience and Remote Sensing* **41**(9): 1907–1919.

Herold, M., Gardner, M., Hadley, B. and Roberts, D. (2002a). The Spectral Dimension in urban land cover mapping from high- resolution optical remote sensing data, *Proceedings of the 3 rd Symposium on Remote Sensing of Urban Areas.*

Herold, M., Scepan, J. and Günther, S. (2002b). Object-oriented mapping and analysis of urban land use/cover using IKONOS data, *Proceedings of the 22nd EARSEL symposium*.

HOT OSM (2015). Humanitarian OpenStreetMap Team, Available from: http://hot.openstreetmap.org/. Last Accessed: 10 December 2014.

Huang, X., Lu, Q., Zhang, L. and Plaza, A. (2014). New postprocessing methods for remote sensing image classification: A systematic study, *IEEE Transactions on Geoscience and Remote Sensing* **52**(11): 7140–7159.

Indirli, M. (2008). Manejo de Riesgos en Valparaíso, Available from: http://www.marvasto.bologna.enea.it/. Last Accessed: 5 March 2015.

Indirli, M., Geremei, F., Puglisi, C. and Screpanti, A. (2008). A GIS platform on main natural hazards for Valparaíso city (Chile) and vulnerability studies for some historical constructions and urban sectors, *in* E. Fodde (ed.), *Structural Analysis of Historic Construction: Preserving Safety and Significance*, Bath, United Kingdom, pp. 1277–1285.

Indirli, M., Valpreda, E., Panza, G., Romanelli, F., Lanzoni, L., Teston, S., Berti, M., Bennardo, S. D. and Rossi, G. (2010). Natural multi-hazard and building vulnerability assessment in the historical centers: the examples of San Giuliano di Puglia (Italy) and Valparaiso (Chile), *Proceedings of the European Commission Conference "SAUVEUR", Safeguarded Cultural Heritage, May 31 - June 3*, University of Trento, Prag.

INE (2015). Instituto Nacional de Estadísticas Chile : Estadísticas Chile, Available from: http://www.ine.cl/canales/chile_estadistico/familias/censos.php. Last Accessed: 22 January 2015.

Johnson, B. and Xie, Z. (2013). Classifying a high resolution image of an urban area using super-object information, *ISPRS Journal of Photogrammetry and Remote Sensing* **83**: 40–49.

Kamel Boulos, M. N., Resch, B., Crowley, D. N., Breslin, J. G., Sohn, G., Burtner, R., Pike, W. A., Jezierski, E. and Chuang, K.-Y. S. (2011). Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples., *International journal of health geographics* **10**: 67.

Kaur, B. and Garg, A. (2011). Mathematical morphological edge detection for remote sensing images, *ICECT 2011 - 2011 3rd International Conference on Electronics Computer Technology*, Vol. 5, pp. 324–327.

Kavzoglu, T. and Colkesen, I. (2013). An assessment of the effectiveness of a rotation forest ensemble for land-use and land-cover mapping, *International Journal of Remote Sensing* **34**(12): 4224–4241.

Kinley, L. (2013). Towards the use of Citizen Sensor Information as an Ancillary Tool for the Thematic Classification of Ecological Phenomena, *Proceedings of the 2nd AGILE (Association of Geographic Information Laboratories for Europe) PhD School 2013 of the 2nd AGILE*.

Klonner, C., Barron, C., Neis, P. and Höfle, B. (2014). Updating digital elevation models via change detection and fusion of human and remote sensor data in urban environments, *International Journal of Digital Earth* **8**(2): 151–169.

Kounadi, O. (2009). *Assessing the quality of OpenStreetMap data*, Masterthesis, University College of London.

Kuncheva, L. I. and Rodriguez, J. J. (2007). An Experimental Study on Rotation Forest Ensembles, *Lecture Notes in Computer Science* **4472**: 459–468.

Kunze, C. and Hecht, R. (2015). Semantic enrichment of building data with volunteered geographic information to improve mappings of dwelling units and population, *Computers, Environment and Urban Systems*.

Landis, J. and Koch, G. (1977). The measurement of observer agreement for categorical data, *Biometrics* **33**: 159 – 174.

Langford, M. (2004). Obtaining population estimates in non-census reporting zones: An evaluation of the 3-class dasymetric method, *Computers, Environments and Urban Systems* **30**(2): 161–180.

Lari, Z. and Ebadi, H. (2007). Automated Building Extraction from High-Resolution Satellite Imagery Using Spectral and Structural Information Based on Artificial Neural Networks, *ISPRS Hannover Workshop*, number 1346.

Lasota, T., Luczak, T. and Trawi, B. (2012). Investigation of Rotation Forest Method Applied to Property Price Prediction, *Artificial Intelligence and Soft Computing* **7267**: 403–411.

Li, P., Guo, J., Song, B. and Xiao, X. (2011). A multilevel hierarchical image segmentation method for urban impervious surface mapping using very high resolution imagery, *Selected Topics in Applied Earth Observations and Remote Sensing (IEEE)* **4**(1): 103–116.

Li, P., Guo, J., Xu, H. and Xiao, X. (2008). Multilevel Object based Image Classification over Urban Area based hierarchical Image Segmentation and Invariant Moments, *Proceedings of GEOBIA*, Calgary Alberta, Canada.

Li, W., Wang, S. and Li, J. (2014). Object Based Building Extraction by QuickBird Image for Population Estimation, *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*, Quebec City, Canada, pp. 3176 – 3179.

Lillesand, T. M., Kiefer, R. W. and Chipman, J. W. (2004). *Remote Sensing and Image Interpretation*, 5 edn, Wiley.

Lim, J. S. (1990). *Two-Dimensional Signal and Image Processing*, 1 edn, Prentice Hall, New Jersey.

Liu, X. H., Kyriakidis, P. C. and Goodchild, M. F. (2008). Population-density estimation using regression and area-to-point residual kriging, *International Journal of Geographical Information Science* **22**(4): 431–447.

Liu, Z., Wang, J. and Liu, W. (2005). Building extraction from high resolution imagery based on multi-scale object oriented classification and probabilistic Hough transform, *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS 05.* **4**(C): 0–3.

Machine Learning Group at the University of Waikato (2015). WEKA, Available from: http://www.cs.waikato.ac.nz/ml/weka/. Last Accessed: 26 March 2015.

Mackay, D. J. (1998). Introduction to Gaussian Processes, Available from: http://www.inference.eng.cam.ac.uk/mackay/gpB.pdf. Last Accessed: 30 April 2015.

Makridakis, S. and Hibon, M. (1995). Evaluating Accuracy (or Error) Measures, R&D / INSEAD, Fontainebleau, France.

Manfré, L. a., Hirata, E., Silva, J. B., Shinohara, E. J., Giannotti, M., Larocca, A. P. C. and Quintanilha, J. (2012). An Analysis of Geospatial Technologies for Risk and Natural Disaster Management, *ISPRS International Journal of Geo-Information* **1**(2): 166–185.

Marconcini, M., Fernández-prieto, D. and Buchholz, T. (2014). Targeted Land-Cover Classification, *IEEE Transactions on Geoscience and Remote Sensing* **52**(7): 4173–4193.

Martha, T. R., Kerle, N., Jetten, V., van Westen, C. J. and Kumar, K. V. (2010). Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods, *Geomorphology* **116**(1-2): 24–36.

Mennis, J. and Hultgren, T. (2006). Intelligent Dasymetric Mapping and Its Application to Areal Interpolation, *Cartography and Geographic Information Science* **33**(3): 179–194.

Mianji, F. a. and Zhang, Y. (2011). Robust hyperspectral classification using relevance vector machine, *IEEE Transactions on Geoscience and Remote Sensing* **49**(6): 2100–2112.

Misakova, L., Jacquin, A. and Gay, M. (2006). Mapping Urban Sprawl Using VHR Data and Object Oriented Classification, *Proceedings of the 1st Workshop of the EARSeL SIG on Urban Remote Sensing*, Berlin, Germany.

Moine, M., Puissant, A. and Malet, J. (2009). Detection of landslides from aerial and satellite images with a semi-automatic method. Application to the Barcelonnette basin (Alpes-de-Hautes-Provence, France), *Landslide Processes: from geomorphological mapping to dynamic modelling*, pp. 63–68.

Mooney, P. and Corcoran, P. (2012). Characteristics of Heavily Edited Objects in OpenStreetMap, *Future Internet* **4**(4): 285–305.

Mooney, P., Corcoran, P. and Winstanley, A. C. (2010). Towards quality metrics for OpenStreetMap, *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ACM Press, New York, USA, pp. 514–517.

Mucke, P., Marx, M., Welle, T., Birkmann, J., Rhyner, J., Witting, M. and Wolfertz, J. (2013). World Risk Report 2013, *Technical report*.

Mueller, M., Segl, K., Heiden, U. and Kaufmann, H. (2006). Potential of high-resolution satellite data in the context of vulnerability of buildings, *Natural Hazards* **38**(1-2): 247–258.

Munivalpo (2015). Municipalidad de Valparaíso - Ciudad Patrimonio de la Humanidad, Available from: http://www.municipalidaddevalparaiso.cl/valparaiso\. Last Accessed: 22 January 2015.

Myint, S. W., Gober, P., Brazel, A., Grossman-Clarke, S. and Weng, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery, *Remote Sensing of Environment* **115**(5): 1145–1161.

Neis, P. (2015). OSMstats, Available from: http://osmstats.neis-one.org/?item=members. Last Accessed: 13 January 2015.

Neis, P. and Zipf, A. (2012). Analyzing the Contributor Activity of a Volunteered Geographic Information Project — The Case of OpenStreetMap, *ISPRS International Journal of Geo-Information* **1**(3): 146–165.

Neis, P., Zielstra, D. and Zipf, A. (2013). Comparison of Volunteered Geographic Information Data Contributions and Community Development for Selected World Regions, *Future Internet* **5**(2): 282–300.

Okujeni, A., van der Linden, S., Tits, L., Somers, B. and Hostert, P. (2013). Support vector regression and synthetically mixed training data for quantifying urban land cover, *Remote Sensing of Environment* **137**: 184–197.

ONEMI (2014). Ministerio de Interior y Seguridad Pública: Incendio en Valparaíso, Available from: http://www.onemi.cl/incendio-en-valparaiso. Last Accessed: 22 January 2015.

OSM (2015a). Bing Image Analyser, Available from: http://ant.dev.openstreetmap.org/bingimageanalyzer/. Last Accessed: 22 January 2015.

OSM (2015b). OpenStreetMap Bing Imagery, Available from: http://mvexel.dev.openstreetmap.org/bing/. Last Accessed: 22 January 2015.

OSM (2015c). Openstreetmap copyright, Available from: http://www.openstreetmap.org/copyright/de. Last Accessed: 13 January 2015.

OSM (2015d). OpenStreetMap Map Features, Available from: http://wiki.openstreetmap.org/wiki/Map_Features. Last Accessed: 13 January 2015.

OSM (2015e). Planet.osm, Available from: http://wiki.openstreetmap.org/wiki/Planet.osm. Last Accessed: 22 January 2015.

OSM–Wiki (2015). Bing, Available from: http://wiki.openstreetmap.org/wiki/Bing. Last Accessed: 22 January 2015.

OSM Task Manager (2014a). #502 - Valparaíso, Chile Fires / Fuegos en Valparaíso, Chile, Available from: http://tasks.hotosm.org/project/502. Last Accessed: 22 January 2015.

OSM Task Manager (2014b). #508 - Valparaíso, Chile Fires 2 / Valparaíso, Chile Incendios, Available from: http://tasks.hotosm.org/project/508. Last Accessed: 22 January 2015.

OSM-Wiki (2015). Imagery offset database, Available from: http://wiki.openstreetmap.org/wiki/Imagery_Offset_Database. Last Accessed: 22 January 2015.

Paladini, R., Martorella, M. and Berizzi, F. (2011). Classification of man-made targets via invariant coherency-matrix eigenvector decomposition of polarimetric SAR/ISAR images, *IEEE Transactions on Geoscience and Remote Sensing* **49**(8): 3022–3034.

Palmason, J., Benediktsson, J. and Arnason, K. (2003). Morphological transformations and feature extraction of urban data with high spectral and spatial resolution, *2003 IEEE International Geoscience and Remote Sensing Symposium*, pp. 7929–7931.

Pasolli, E. and Melgani, F. (2011). Gaussian process regression within an active learning scheme, *2011 IEEE International Geoscience and Remote Sensing Symposium* pp. 3574–3577.

Pasolli, L., Melgani, F. and Blanzieri, E. (2008). Estimating Biophysical Parameters from Remotely Sensed Imagery with Gaussian Processes, *IGARSS 2008 - IEEE International Geoscience and Remote Sensing Symposium* **2**: 851–854.

Peijun, L., Hongtao, H. and Jiancong, G. (2007). Segmentation of high-resolution multispectral image based on extended morphological profiles, *International Geoscience and Remote Sensing Symposium (IGARSS)* pp. 1481–1484.

Pesaresi, M. and Benediktsson, J. (2001). A new approach for the morphological segmentation of high-resolution satellite imagery, *IEEE Transactions on Geoscience and Remote Sensing* **39**(2): 309–320.

Pesaresi, M., Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., Halkia, M., Kauffmann, M., Kemper, T., Lu, L., Marin-herrera, M. A., Ouzounis, G. K., Scavazzon, M., Soille, P., Syrris, V. and Zanchetta, L. (2013). A Global Human Settlement Layer From Optical HR/VHR RS Data : Concept and First Results, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **6**(5): 2102–2131.

Pierre Soille (2003). *Morphological Image Analysis. Principles and Applications*, 2 edn, Springer Press, Berlin.

Pina, P. and Barata, T. (2003). Classification by mathematical morphology, *2003 IEEE International Geoscience and Remote Sensing Symposium* **6**(C): 3516–3518.

Pittore, M., Wieland, M. and Fleming, K. (2015). Perspectives of a global, dynamic Exposure Model for Geo-Risk Assessment from Remote Sensing to Crowd-Sourcing, *Global Assessment Report on Disaster Risk Reduction 2015*.

Poser, K. and Dransch, D. (2010). Volunteered Geographic Information for Disaster Management With Application To Rapid Flood Damage Estimation, *Geomatica* **64**(1): 89–98.

Preventionweb (2015). Chile - Disaster Statistics, Available from: http://www.preventionweb.net/english/countries/statistics/?cid=35. Last Accessed: 17 February.

Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*, MIT Press.

Res, C., Willmott, C. J. and Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, *Climate Research (Clim Res)* **30**: 79–82.

Richter, R. (1996). A spatially adaptive fast atmospheric correction algorithm, *International Journal of Remote Sensing* **17**(6): 1201–1214.

Richter, R. and Schläpfer, D. (2014). Atmospheric/Topographic Correction for Satellite Imagery, *Technical report*.

Rizvi, I. A. and Mohan, B. K. (2010). Improving the Accuracy of Object Based Supervised Image Classification using Cloud Basis Function Neural Network for High Resolution Satellite Images, *Image Processing* (4): 342–353.

Rizvi, I. A. and Mohan, B. K. (2012). Object-Based Analysis of Worldview-2 Imagery of Urban Areas, *IEEE Transactions on Geoscience and Remote Sensing* **6**: 431–434.

Rodriguez, J. J., Kuncheva, L. I. and Alonso, C. J. (2006). Rotation Forest: A New Classifier Ensemble Method, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**(10): 1619–1630.

Rokach, L. and Maimon, O. (2008). *Data mining with decision trees: theory and applications*, World Scientific Pub Co Inc.

Römer, H., Willroth, P., Kaiser, G., Vafeidis, a. T., Ludwig, R., Sterr, H. and Revilla Diez, J. (2012). Potential of remote sensing techniques for tsunami hazard and vulnerability analysis – a case study from Phang-Nga province, Thailand, *Natural Hazards and Earth System Science* **12**(6): 2103–2126.

Samaniego, L., Bárdossy, A. and Schulz, K. (2008). Supervised Classification of Remotely Sensed Imagery Using a Modified k-NN Technique, *IEEE Transactions on Geoscience and Remote Sensing* **46**(7): 1–26.

Sánchez M., A., Bosque M., J. and Jiménez V., C. (2009). Valparaíso: su geografía, su historia y su identidad como Patrimonio de la Humanidad, *Estudios Geográficos* **LXX**(266): 269–293.

Schnebele, E. and Cervone, G. (2013). Improving remote sensing flood assessment using volunteered geographical data, *Natural Hazards and Earth System Science* **13**(3): 669–677.

Sester, M., Arsanjani, J. J., Klammer, R., Burghardt, D. and Haunert, J.-h. (2014). Integrating and Generalizing Volunteered Geographic Information, *in* W. Burghardt, B., Duchene, C., Mackaness (ed.), *Abstracting Geographic Information in a Data Rich World*, Springer Press.

Shackelford, A. K. and Davis, C. H. (2003). A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas, *IEEE Transactions on Geoscience and Remote Sensing* **41**(10): 2354–2363.

Sheng, L. S. L., Xiaoyu, W. X. W., Xinfa, Q. X. Q. and Yongjian, H. Y. H. (2009). Mathematical Morphology Edge Detection Algorithm of Remote Sensing Image with High Resolution, *Information Science and Engineering (ICISE), 2009 1st International Conference on* (4): 1323–1326.

SHOA (2015). Chilenian Navy Hydrographic and Oceanographic Service, Available from: www.shoa.cl/index.htm. Last Accessed: 07 May 2015.

SIC (2015). Satellite Imaging Cooperation: Worldview-2 Satellite Sensor, Available from: http://www.satimagingcorp.com/satellite-sensors/worldview-2/. Last Accessed: 30 March 2015.

Singh, D., Maurya, R., Shukla, A. S., Sharma, M. K. and Gupta, P. R. (2012). Building extraction from very high resolution multispectral images using NDVI based segmentation and morphological operators, *2012 Students Conference on Engineering and Systems*, IEEE, pp. 1–5.

Sirmacek, B. and Unsalan, C. (2011). A probabilistic framework to detect buildings in aerial and satellite images, *IEEE Transactions on Geoscience and Remote Sensing* **49**(1): 211–221.

Smola, A. J. and Schölkopf, B. (2004). A tutorial on support vector regression, *Statistics and Computing* **14**: 199–222.

Soille, P. and Pesaresi, M. (2002). Advances in mathematical morphology applied to geoscience and remote sensing, *IEEE Transactions on Geoscience and Remote Sensing* **40**(9): 2042–2055.

Steinnocher, K., Köstl, M. and Weichselbaum, J. (2011). Kleinräumige Bevölkerungsmodellierung für Europa – räumliche Disaggregation auf Basis des Versiegelungsgrades, *in* J. Strobl, T. Blaschke and G. Griesebner (eds), *Angewandte Geoinformatik 2011*, Wichmann, Berlin, pp. 513–520.

Stiglic, G., Rodriguez, J. J. and Kokol, P. (2011). Rotation of Random Forests for genomic and proteomic classification problems, *Advances in Experimental Medicine and Biology*, Vol. 696, pp. 211–221.

Stumpf, A. and Kerle, N. (2011). Object-oriented mapping of landslides using Random Forests, *Remote Sensing of Environment* **115**(10): 2564–2577.

Su, M.-D., Lin, M.-C., Hsieh, H.-I., Tsai, B.-W. and Lin, C.-H. (2010). Multi-layer multi-class dasymetric mapping to estimate population distribution., *The Science of the total environment* **408**(20): 4807–16.

Syed, S., Dare, P. and Jones, S. (2005). Automatic Classification of Land Cover Features with High Resolution Imagery and LIDAR Data : An Object-Oriented Approach, *Proceedings of the National Biennial Conference of the Spatial Sciences Institute*, Melbourne, Australia, pp. 512–522.

Taubenböck, H. (2008). A conceptual vulnerability and risk framework as outline to identify capabilities of remote sensing, *Natural Hazards and Earth System Science* **8**: 409–420.

Taubenböck, H., Goseberg, N., Lämmel, G., Setiadi, N., Schlurmann, T., Nagel, K., Siegert, F., Birkmann, J., Traub, K. P., Dech, S., Keuck, V., Lehmann, F., Strunz, G. and Klüpfel, H. (2013). Risk reduction at the "Last-Mile": An attempt to turn science into action by the example of Padang, Indonesia, *Natural Hazards* **65**(1): 915–945.

Taubenböck, H., Roth, A. and Dech, S. (2007). Vulnerability assessment using remote sensing: The earthquake prone mega-city Istanbul, Turkey, *Proceedings of ISRSE 2007 , 32nd International Symposium on Remote Sensing on Environment, June 25 - June 27*, San Jose, Costa Rica.

Taubenböck, H., Wurm, M., Netzband, M., Zwenzner, H., Roth, A., Rahman, A. and Dech, S. (2011). Flood risks in urbanized areas - Multi-sensoral approaches using remotely sensed data for risk assessment, *Natural Hazards and Earth System Science* **11**(2): 431–444.

Taubenböck, H., Wurm, M., Setiadi, N., Gebert, N., Roth, A., Strunz, G., Birkmann, J. and Dech, S. (2009). Integrating Remote Sensing and Social Science - The correlation of urban morphology with socioeconomic parameters, *Urban Remote Sensing Joint Event, May 20-22*, Shanghai, China.

The World Bank Group (2015). Population Statistics, Available from: http://data.worldbank.org/indicator/SP.POP.TOTL. Last Accessed: 07 May 2015.

Thomas, N., Hendrix, C. and Congalton, R. G. (2003). A Comparison of Urban Mapping Methods Using High-Resolution Digital Imagery, *Photogrammetric Engineering and Remote Sensing* **69**(9): 963–972.

Trimble GmbH (2014). *Definiens Developer 7 Reference Book*, Trimble Germany GmbH.

Tuia, D. and Camps-Valls, G. (2009). Cluster kernels for semisupervised classification of VHR urban images, *2009 Joint Urban Remote Sensing Event*.

Tuia, D. and Camps-Valls, G. (2011). Urban image classification with semisupervised multiscale cluster kernels, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **4**(1): 65–74.

Tuia, D., Pacifici, F., Kanevski, M. and Emery, W. J. (2009). Classification of very high spatial resolution imagery using mathematical morphology and support vector machines, *IEEE Transactions on Geoscience and Remote Sensing* **47**(11): 3866–3879.

Tuia, D., Verrelst, J., Alonso, L., Perez-Cruz, F. and Camps-Valls, G. (2011). Multioutput support vector regression for remote sensing biophysical parameter estimation, *IEEE Geoscience and Remote Sensing Letters* **8**(4): 804–808.

Turker, M. and Koc San, D. (2010). Building Detection from Pan-Sharpened IKONOS Imagery through Support Vector Machines Clasiification, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS)* **XXXVIII**(8): 841–846.

Turner, B. L., Kasperson, R. E., Matson, P. a., McCarthy, J. J., Corell, R. W., Christensen, L., Eckley, N., Kasperson, J. X., Luers, A., Martello, M. L., Polsky, C., Pulsipher, A. and Schiller, A. (2003). A framework for vulnerability analysis in sustainability science, *Proceedings of the National Academy of Sciences of the United States of America* **100**(14): 8074–8079.

UN (2003). United Nations World Heritage Convention: World Heritage List - Historic Quarter of the Seaport City of Valparaíso, Available from: http://whc.unesco.org/en/list/959. Last Accessed: 23 April 2015.

UNISDR (2004). *Living with risk: a global review of disaster reduction initiatives*, Vol. 1, United Nations Inter-Agency Secretariat of the International Strategy for Disaster Reduction (UN/ISDR).

UNISDR (2015). International Strategy of Disaster Risk Reduction, Available from: http://www.unisdr.org/. Last Accessed: 23 April 2015.

Van de Voorde, T., De Genst, W. and Canters, F. (2007). Improving pixel-based VHR land-cover classifications of urban areas with post-classification techniques, *Photogrammetric Engineering and Remote Sensing* **73**(9): 1017–1027.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*, Springer Press, New York.

Verrelst, J., Muñoz, J., Alonso, L., Delegido, J., Rivera, J. P., Camps-Valls, G. and Moreno, J. (2012). Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3, *Remote Sensing of Environment* **118**: 127–139.

Wang, Y. and Witten, I. H. (1997). Induction of model trees for predicting continuous classes, *Proceedings of the 9th European Conference on Machine Learning Poster Papers*, pp. 128–137.

Waske, B., Fauvel, M., Benediktsson, J. A. and Chanussot, J. (2009). Machine learning techniques in remote sensing data analysis, *in* G. Camps-valls and L. Bruzzone (eds), *Kernel Methods for Remote Sensing Data Analysis*, John Wiley and Sons, pp. 1–24.

Wei, Y., Zhao, Z. and Song, J. (2008). Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection, *IEEE Transactions on Geoscience and Remote Sensing* **00**(c): 2008–2010.

WEKA Data Mining (2015). Forecasting evaluation models, Available from: http://weka.sourceforge.net/doc.packages/timeseriesForecasting/weka/classifiers/timeseries/eval/package-summary.html. Last Accessed: 07 May 2015.

Weng, Q. and Hu, X. (2008). Medium spatial resolution satellite imagery for estimating and mapping urban impervious surfaces using LSMA and ANN, *IEEE Transactions on Geoscience and Remote Sensing* **46**(8): 2397–2406.

Weng, Q. and Quattrochi, D. D. (2008). *Urban Remote Sensing*, CRC Press, Boca Raton.

White, P., Pelling, M., Sen, K., Seddon, D., Russell, S. and Few, R. (2005). Disaster risk reduction. A development concern, *DFID Department for International Development* pp. 1–8.

Wieland, M., Pittoire, M., Parolai, S., Zschau, J., Moldobekov, B. and Begaliev, U. (2012a). Estimating building inventory for rapid seismic vulnerability assessment: Towards an integrated approach based on multi-sourcing imaging, *Soil dynamics and Earthquake engineering* **36**: 70–83.

Wieland, M., Pittore, M., Parolai, S. and Zschau, J. (2012b). Estimating building inventory for rapid seismic vulnerability assessment in Bishkek , Kyrgyzstan, *Soil dynamics and Earthquake engineering* **36**: 70–83.

Wieland, M., Pittore, M., Parolai, S. and Zschau, J. (2012c). Exposure Estimation from Multi-Resolution Optical Satellite Imagery for Seismic Risk Assessment, *ISPRS International Journal of Geo-Information* **1**(3): 69–88.

Wiesner, M. and Taubenböck, H. (2014). It's Not Big, It's Large: Mapping and Characterizing Urban Landscapes of a Different Magnitude based on EO Data, *in* M. Schrenk, V. V. Popvich, P. Zeile and P. Elisei (eds), *Proceedings REAL CORP, May 21-23*, Vol. 8, pp. 319–328.

Wisner, B., Blaikie, P., Cannon, T. and Davis, I. (2003). *At Risk : natural hazards , people ' s vulnerability and disasters*, Routledge Chapman & Hall.

Wu, S.-S., Qiu, X. and Wang, L. (2005). Population Estimation Methods in GIS and Remote Sensing: A Review, *GIScience and Remote Sensing* **42**(1): 58–74.

Wurm, M., Taubenböck, H., Schardt, M., Esch, T. and Dech, S. (2011). Object-based image information fusion using multisensor earth observation data over urban areas, *International Journal of Image and Data Fusion* **2**(2): 121–147.

Xia, J., Member, S., Du, P., Member, S., He, X. and Chanussot, J. (2014). Hyperspectral Remote Sensing Image Classification Based on Rotation Forest, *IEEE Geoscience and Remote Sensing Letters* **11**(1): 239–243.

Zheng, L., Wan, L., Huo, H. and Fang, T. (2014). A noise removal approach for object-based classification of VHR imagery via post-classification, *2014 International Conference on Audio, Language and Image Processing*, pp. 915–920.

Zhong, Y. and Zhang, L. (2012). An Adaptive Artificial Immune Network for Supervised Classification of Multi-/Hyperspectral Remote Sensing Imagery, *IEEE Transactions on Geoscience and Remote Sensing* **50**(3): 894–909.

Zhou, Z.-h. (2009). Ensemble Learning, *in* S. Z. Li (ed.), *Encyclopedia of Biometrics*, Springer, Berlin, pp. 270–273.

Zielstra, D. and Zipf, A. (2010). A Comparative Study of Proprietary Geodata and Volunteered Geographic Information for Germany, *Proceedings of 13th AGILE International Conference on Geographic Information Science, May 10-14*, Vol. 1, Guimarães, Portugal.

A. Appendix

A.1. Tiling for Processing

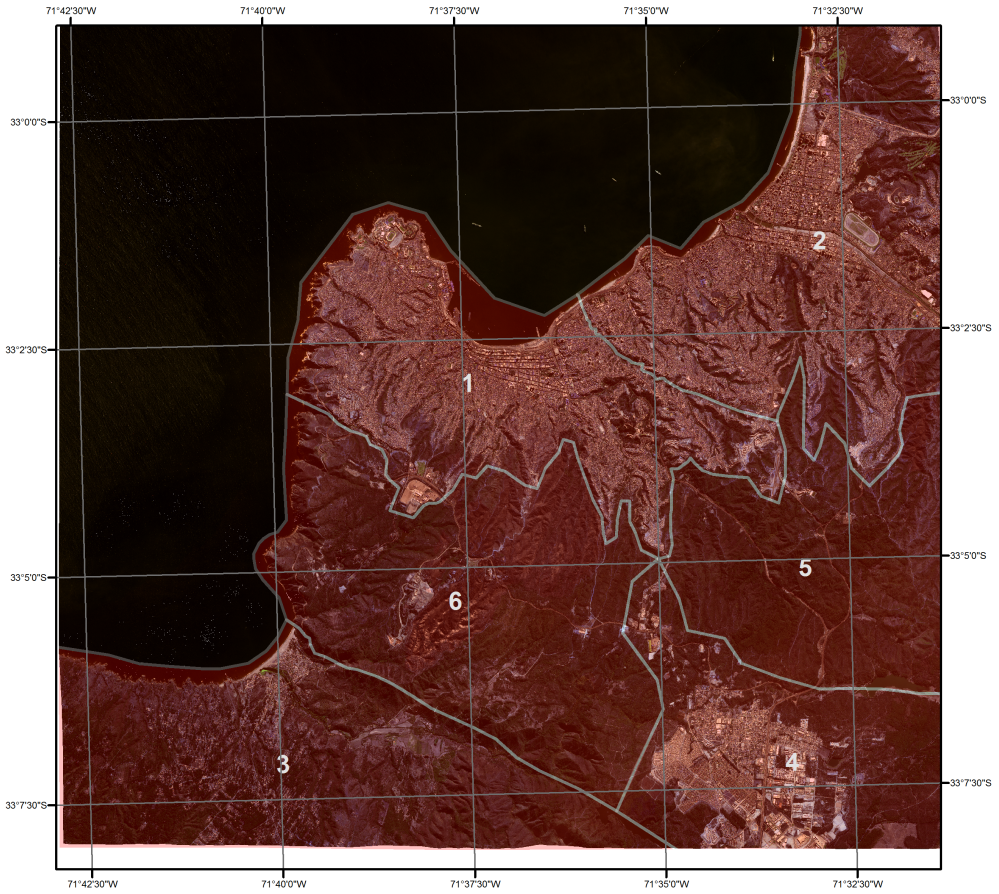


Figure A.1.: Tiling for processing: In terms of processing times, the scene is divided into 6 subsets in accordance to natural transition of landcover.

A.2. Categories of Ancillary OSM Information

Category	key;value
Industrial	landuse; industrial building; industry
Culture	amenity; cinema, theater tourism;attraction, museum, artwork historic;heritage, monument
Leisure	landuse; park, recreation_ground, villagegreen leisure;garden,park,pitch, playground, sportcentre, stadium, swimming-pool,track tourism;viewpoint amenity;swimming_pool building; stadium
Gastronomy	amenity;bar, cafe, fast food, pub, restaurant
Education	buidling; university, school amenity; university, college, school, kindergarden
Traffic	building; transportation amenity; bus station, parking, taxi highways; bus stop, traffic signals, crossing railway; level crossing
Health	building; hospital amenity; pharmacy, hospital, doctors, dentist, veterinary
Tourism	tourism;guest house, hotel, hostel, motel building; hotel
Services	shop; hairdresser, laundry, travel agency amenity; post office, toilets, library, atm, bank
Shopping	shop; convenience, department store, doityourself, medical supply, radiotechnics, supermarket, bakery, alcohol, beverages, butcher, confectionery, convenience, seafood, kiosk, mall, shoes, clothes amenity; marketplace buidling; retail, commercial,
Religion	amenity; place of worship, graveyard landuse;cementary buidling; church
Public	amenity; community center, courthouse, police, prison, townhall, fire station building; public

Table A.1.: Categories of ancillary OSM information: OSM landuse information and POIs are classified in 12 predefined categories.

A.3. Entire LULC Map

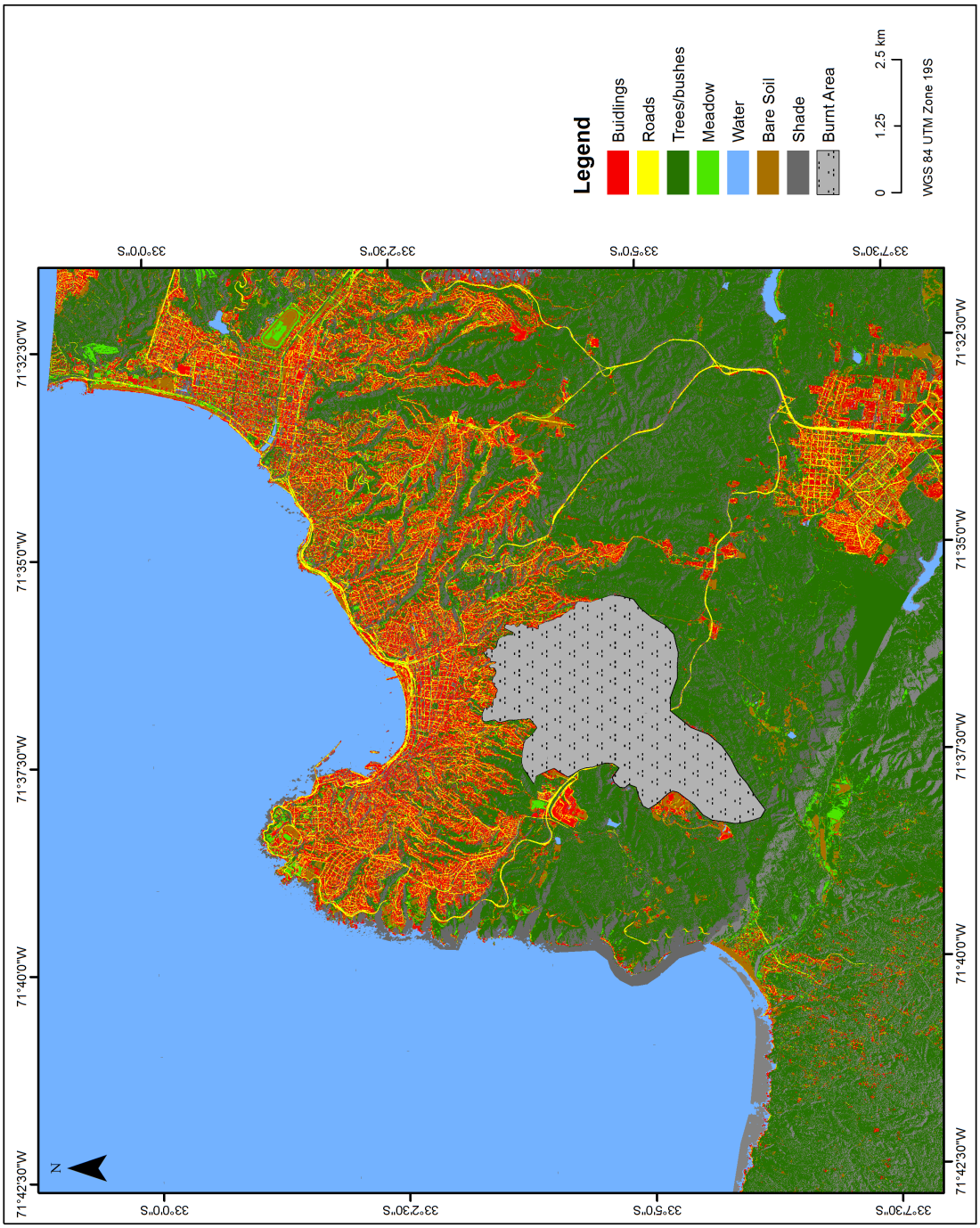


Figure A.2.: LULC Map after for the entire Worldview-2 scene: Result of classification by means of RTF classifier and classification postprocesing